

Cooperative Transmission for a Vector Gaussian Parallel Relay Network

Muryong Kim, *Member, IEEE*, and Sae-Young Chung, *Senior Member, IEEE*

Abstract

In this paper, we consider a parallel relay network where two relays cooperatively help a source transmit to a destination. We assume the source and the destination nodes are equipped with multiple antennas. Three basic schemes and their achievable rates are studied: Decode-and-Forward (DF), Amplify-and-Forward (AF), and Compress-and-Forward (CF). For the DF scheme, the source transmits two private signals, one for each relay, where dirty paper coding (DPC) is used between the two private streams, and a common signal for both relays. The relays make efficient use of the common information to introduce a proper amount of correlation in the transmission to the destination. We show that the DF scheme achieves the capacity under certain conditions. We also show that the CF scheme is asymptotically optimal in the high relay power limit, regardless of channel ranks. It turns out that the AF scheme also achieves the asymptotic optimality but only when the relays-to-destination channel is full rank. The relative advantages of the three schemes are discussed with numerical results.

Index Terms

Gaussian parallel relay network, diamond channel, cooperative relaying, common information.

I. INTRODUCTION

Over the recent years, relaying has been considered as a promising technique that can increase throughput and reliability and enhance the coverage of wireless networks. There have been a

This work was supported in part by the Center for Broadband OFDM Mobile Access (BrOMA) through the ITRC Program of the Ministry of Information and Communication (MIC), Korea, supervised by the Institute of Information Technology Assessment (IITA). The material in this paper was presented in part at the 45th Annual Allerton Conference on Communication, Control, and Computing, Monticello, IL, September 2007.

The authors are with the School of Electrical Engineering and Computer Science, KAIST, Daejeon, Korea (e-mail: muryong@kaist.ac.kr, sychung@ee.kaist.ac.kr).

number of research results showing different aspects of relay channels. Cover and El Gamal [3] derived the capacity of a class of relay channels with a single relay that helps transmission from a source to a destination. Kramer et al. [4] generalized the results in various ways. Laneman et al. [5] considered cooperative diversity aspects of relay channels. In [23], the authors consider the multiplexing aspects of cooperative communications in multi-antenna relay networks. In [6], [7], multi-antenna relay channels are considered. The use of relays in broadcast scenarios is considered in [4], [8].

In this paper, we study the capacity of a vector Gaussian parallel relay network where two parallel relays help transmission from a source node to a destination node. The network model is first studied by Schein et al. [1], and a set of capacity theorems are derived for the discrete memoryless channel and the scalar Gaussian channel. The authors of [2] considered a similar model but with half-duplex constraint, i.e., relays do not transmit and receive at the same time. Recently, the authors of [22] showed their new achievable rate for general scalar Gaussian relay networks is within a constant number of bits from the cut-set upper bound on the capacity. A new achievable rate for the original Schein's network is derived using a Combined Amplify-and-Decode Forward (CADF) scheme in [24]. Our network model is different from the earlier ones in that both the broadcast channel (BC) part and the multiple access channel (MAC) part are vector Gaussian channels as the source and the destination nodes are equipped with multiple antennas. As the vector BC is not degraded in general and a simple superposition coding will not suffice. In the vector MAC, correlation between relay signals are not always beneficial, rather, the right amount of correlation may result in a better performance as will be seen in a later section. Throughout the paper, upper bounds and achievable rates by different cooperative transmission strategies: DF, AF and CF are derived.

For the DF scheme, the vector Gaussian parallel relay network can be seen as a cascade of multiple-input single-output (MISO) BC and single-input multiple-output (SIMO) MAC channels. We first extend some earlier results for the discrete memoryless and scalar MACs with common information to the vector Gaussian case, and investigate the characteristics of the three-dimensional achievable rate region. We use a known transmission scheme of [10] for the BC part. Using the BC-MAC schemes, we show that DF achieves the capacity of the vector Gaussian parallel relay network under certain conditions. In addition, we address the importance of common information signaling and correlation control.

We also extend some earlier results for AF and CF to the vector Gaussian case. We show when DF is strictly suboptimal and when AF and CF can outperform DF by comparing the achievable rates and the upper bound. We show that AF is asymptotically optimal in the high relay power limit if the channel rank of the MAC part is full. In addition, we also show that the CF scheme achieves the asymptotic capacity, regardless of the channel ranks.

The rest of the paper is organized as follows. In Section II, we introduce the system model. We derive a capacity upper bound for the vector Gaussian parallel relay network in Section III. Next, we derive achievable rates by the DF, AF and CF schemes in Sections IV, V and VI, respectively. Numerical results and comparison of different schemes are given in Section VII. Conclusions and final remarks are given in Section VIII.

II. SYSTEM MODEL

The vector Gaussian parallel relay network consists of four nodes: a source, a destination, and two relays. We assume no direct link from the source to the destination. The relays are assumed to be full duplex, i.e., they transmit and receive at the same time. The received signals at the relays and at the destination are given by

$$\begin{aligned} y_{r1} &= \mathbf{g}_1 \mathbf{x}_s + n_{r1}, \\ y_{r2} &= \mathbf{g}_2 \mathbf{x}_s + n_{r2}, \\ \mathbf{y}_d &= \mathbf{h}_1 x_{r1} + \mathbf{h}_2 x_{r2} + \mathbf{n}_d, \end{aligned}$$

where

- $\mathbf{x}_s \in \mathbb{C}^{M \times 1}$, $x_{r1}, x_{r2} \in \mathbb{C}$ are the transmitted signals from the source and from relays 1 and 2, respectively. Input covariance matrix and power constraint at the source node are given by $\mathbf{Q}_s = \mathbb{E}[\mathbf{x}_s \mathbf{x}_s^\dagger]$ and $\text{tr}(\mathbf{Q}_s) \leq P_s$, respectively. Power constraints at relays are given by $\mathbb{E}[|x_{r1}|^2] \leq P_{r1}$ and $\mathbb{E}[|x_{r2}|^2] \leq P_{r2}$;
- $y_{r1}, y_{r2} \in \mathbb{C}$, $\mathbf{y}_d \in \mathbb{C}^{N \times 1}$ are the received signals at relays 1 and 2 and at the destination, respectively;
- $\mathbf{g}_1, \mathbf{g}_2 \in \mathbb{C}^{1 \times M}$ are the channel gains from the source to relays 1 and 2, respectively, and $\mathbf{h}_1, \mathbf{h}_2 \in \mathbb{C}^{N \times 1}$ are the channel gains from relays 1 and 2 to the destination, respectively;
- $n_{r1}, n_{r2} \in \mathbb{C}$, $\mathbf{n}_d \in \mathbb{C}^{N \times 1}$ are additive white Gaussian noise (AWGN) at relays 1 and 2 and at the destination, respectively. Noise at the relays and each antenna of the destination node

is circularly symmetric complex Gaussian, i.e., $n_{r1}, n_{r2} \sim \mathcal{CN}(0, 1)$ and $\mathbf{n}_d \sim \mathcal{CN}(0, \mathbf{I})$ and they are all independent of each other and from the signals.

Throughout the paper, the following notation will be used,

- The vector and matrix notations:

$$\mathbf{y}_r = \begin{bmatrix} y_{r1} \\ y_{r2} \end{bmatrix}, \mathbf{x}_r = \begin{bmatrix} x_{r1} \\ x_{r2} \end{bmatrix}, \mathbf{n}_r = \begin{bmatrix} n_{r1} \\ n_{r2} \end{bmatrix}, \mathbf{G} = \begin{bmatrix} \mathbf{g}_1 \\ \mathbf{g}_2 \end{bmatrix}, \mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 \end{bmatrix}.$$

- $\mathbf{x}_s^n, x_{r1}^n, x_{r2}^n, y_{r1}^n, y_{r2}^n, \mathbf{y}_d^n$ denote length- n sequences of $\mathbf{x}_s, x_{r1}, x_{r2}, y_{r1}, y_{r2}, \mathbf{y}_d$, respectively.
- $\mathbb{E}_{\mathbf{x}}[\cdot]$ denotes expectation with respect to the distribution of \mathbf{x} , $\mathbb{E}_x[\cdot|y]$ does expectation with respect to the distribution of x conditioned on y , and the simpler notation $\mathbb{E}[\cdot]$ without subscript will be used as long as it is apparent.
- $(\cdot)^{opt}$ means the optimal value of a variable.

An example for $M = N = 2$ case where \mathbf{H} and \mathbf{G} are two-input two-output channels is shown in Fig. 1.

Definition 1: A $(2^{nR}, n)$ code for vector Gaussian parallel relay network consists of a message set $W = \{1, 2, \dots, 2^{nR}\}$, an encoding function at the source $f_s : w \in \{1, 2, \dots, 2^{nR}\} \rightarrow \mathbb{C}^{M \times n}$, relaying functions at two relays, $f_{r1,i} : \mathbb{C}^{i-1} \rightarrow \mathbb{C}$ and $f_{r2,i} : \mathbb{C}^{i-1} \rightarrow \mathbb{C}$, respectively, where $1 \leq i \leq n$ is the time index¹ and a decoding function at the destination $g_d : \mathbb{C}^{N \times n} \rightarrow \hat{w} \in \{1, 2, \dots, 2^{nR}\}$. If the message $w \in W$ is sent, the conditional probability of error is defined as $\lambda(w) = \Pr\{g_d(\mathbf{y}_d^n) \neq w | w \text{ sent}\}$. The average probability of error is defined as $P_e^{(n)} = \frac{1}{2^{nR}} \sum_w \lambda(w)$.

Definition 2: If there exists a sequence of $(2^{nR}, n)$ codes with $P_e^{(n)} \rightarrow 0$, the rate R is said to be achievable.

Definition 3: The capacity C of a vector Gaussian parallel relay network is the supremum of the set of achievable rates.

III. CAPACITY UPPER BOUND

In this section, we derive the cut-set bound [21] applied to the vector Gaussian parallel relay network. From the four cuts shown in Fig. 2, we get the following cut-set bound:

$$\max_{p(\mathbf{x}_s, x_{r1}, x_{r2})} \min [I(\mathbf{x}_s; y_{r1}, y_{r2}, \mathbf{y}_d | x_{r1}, x_{r2}), I(x_{r1}, x_{r2}; \mathbf{y}_d), I(\mathbf{x}_s, x_{r1}; \mathbf{y}_d, y_{r2} | x_{r2}), I(\mathbf{x}_s, x_{r2}; \mathbf{y}_d, y_{r1} | x_{r1})].$$

¹This means that the output of a relay at time i depends on all past received symbols.

Using the Markovity of the channel, i.e., $\mathbf{x}_s \leftrightarrow (x_{r1}, x_{r2}) \leftrightarrow \mathbf{y}_d$ and $(x_{r1}, x_{r2}, \mathbf{y}_d) \leftrightarrow \mathbf{x}_s \leftrightarrow (y_{r1}, y_{r2})$, it is easy to get the following loosened cut-set bound:

$$\max_{p(\mathbf{x}_s, x_{r1}, x_{r2})} \min [I(\mathbf{x}_s; y_{r1}, y_{r2}), I(x_{r1}, x_{r2}; \mathbf{y}_d), I(\mathbf{x}_s; y_{r2}) + I(x_{r1}; \mathbf{y}_d|x_{r2}), I(\mathbf{x}_s; y_{r1}) + I(x_{r2}; \mathbf{y}_d|x_{r1})].$$

For example, for the first term we get

$$\begin{aligned} I(\mathbf{x}_s; y_{r1}, y_{r2}, \mathbf{y}_d|x_{r1}, x_{r2}) &= I(\mathbf{x}_s; y_{r1}, y_{r2}|x_{r1}, x_{r2}) + I(\mathbf{x}_s; \mathbf{y}_d|x_{r1}, x_{r2}, y_{r1}, y_{r2}) \\ &\leq I(\mathbf{x}_s; y_{r1}, y_{r2}) + I(\mathbf{x}_s; \mathbf{y}_d|x_{r1}, x_{r2}, y_{r1}, y_{r2}) \\ &= I(\mathbf{x}_s; y_{r1}, y_{r2}) \end{aligned}$$

where the inequality and the last equality follows from the Markovity $(x_{r1}, x_{r2}, \mathbf{y}_d) \leftrightarrow \mathbf{x}_s \leftrightarrow (y_{r1}, y_{r2})$. For the third term we get

$$\begin{aligned} I(\mathbf{x}_s, x_{r1}; \mathbf{y}_d, y_{r2}|x_{r2}) &= I(\mathbf{x}_s, x_{r1}; \mathbf{y}_d|x_{r2}) + I(\mathbf{x}_s, x_{r1}; y_{r2}|x_{r2}, \mathbf{y}_d) \\ &= I(x_{r1}; \mathbf{y}_d|x_{r2}) + I(\mathbf{x}_s; \mathbf{y}_d|x_{r1}, x_{r2}) + I(\mathbf{x}_s, x_{r1}; y_{r2}|x_{r2}, \mathbf{y}_d) \\ &= I(x_{r1}; \mathbf{y}_d|x_{r2}) + I(\mathbf{x}_s, x_{r1}; y_{r2}|x_{r2}, \mathbf{y}_d) \\ &= I(x_{r1}; \mathbf{y}_d|x_{r2}) + I(\mathbf{x}_s; y_{r2}|x_{r2}, \mathbf{y}_d) + I(x_{r1}; y_{r2}|\mathbf{x}_s, x_{r2}, \mathbf{y}_d) \\ &= I(x_{r1}; \mathbf{y}_d|x_{r2}) + I(\mathbf{x}_s; y_{r2}|x_{r2}, \mathbf{y}_d) \\ &\leq I(x_{r1}; \mathbf{y}_d|x_{r2}) + I(\mathbf{x}_s; y_{r2}) \end{aligned}$$

where the third equality follows from the Markovity $\mathbf{x}_s \leftrightarrow (x_{r1}, x_{r2}) \leftrightarrow \mathbf{y}_d$ and the fifth equality and the inequality follow from the Markovity $(x_{r1}, x_{r2}, \mathbf{y}_d) \leftrightarrow \mathbf{x}_s \leftrightarrow (y_{r1}, y_{r2})$.

Note that this cut-set bound is optimized over the joint input distribution $p(\mathbf{x}_s, x_{r1}, x_{r2})$. Using this, we get the following capacity upper bound for the vector Gaussian parallel relay network.

Theorem 1: The capacity of the vector Gaussian parallel relay network is upper bounded by the minimum of the three expressions given by

$$C \leq R_{sum, BC}^{upper} = \max_{\mathbf{Q}_s, \text{tr}(\mathbf{Q}_s) \leq P_s} \log \det (\mathbf{I} + \mathbf{G} \mathbf{Q}_s \mathbf{G}^\dagger), \quad (1)$$

$$C \leq \max_{|\rho| \in [0, 1]} \min [\log \det (\mathbf{I} + \mathbf{H} \mathbf{Q}_r \mathbf{H}^\dagger), \log [(1 + P_s \|\mathbf{g}_2\|^2)(1 + (1 - |\rho|^2) P_{r1} \|\mathbf{h}_1\|^2)]], \quad (2)$$

$$C \leq \max_{|\rho| \in [0, 1]} \min [\log \det (\mathbf{I} + \mathbf{H} \mathbf{Q}_r \mathbf{H}^\dagger), \log [(1 + P_s \|\mathbf{g}_1\|^2)(1 + (1 - |\rho|^2) P_{r2} \|\mathbf{h}_2\|^2)]]. \quad (3)$$

$$\text{where } \mathbf{Q}_r = \mathbb{E}[\mathbf{x}_r \mathbf{x}_r^\dagger] = \begin{bmatrix} P_{r1} & |\rho| e^{j\angle \mathbf{h}_1^\dagger \mathbf{h}_2} \sqrt{P_{r1} P_{r2}} \\ |\rho| e^{-j\angle \mathbf{h}_1^\dagger \mathbf{h}_2} \sqrt{P_{r1} P_{r2}} & P_{r2} \end{bmatrix}.$$

Proof: From the first cut, the following upper bound is derived:

$$\begin{aligned} I(\mathbf{x}_s; y_{r1}, y_{r2}) &= h(y_{r1}, y_{r2}) - h(y_{r1}, y_{r2} | \mathbf{x}_s) \\ &= h(\mathbf{y}_r) - h(\mathbf{n}_r) \\ &\leq \log(2\pi e)^2 \det \mathbb{E}[\mathbf{y}_r \mathbf{y}_r^\dagger] - \log(2\pi e)^2 \det \mathbb{E}[\mathbf{n}_r \mathbf{n}_r^\dagger] \\ &= \log \det \mathbb{E}[\mathbf{n}_r \mathbf{n}_r^\dagger + \mathbf{G} \mathbf{x}_s \mathbf{x}_s^\dagger \mathbf{G}^\dagger] \\ &= \log \det(\mathbf{I} + \mathbf{G} \mathbf{Q}_s \mathbf{G}^\dagger) \end{aligned}$$

where the inequality follows from the fact that the circularly symmetric complex Gaussian maximizes the entropy [19]. From the second cut, the following upper bound is derived:

$$\begin{aligned} I(x_{r1}, x_{r2}; \mathbf{y}_d) &= h(\mathbf{y}_d) - h(\mathbf{y}_d | x_{r1}, x_{r2}) \\ &= h(\mathbf{y}_d) - h(\mathbf{n}_d) \\ &\leq \log(2\pi e)^N \det \mathbb{E}[\mathbf{y}_d \mathbf{y}_d^\dagger] - \log(2\pi e)^N \det \mathbb{E}[\mathbf{n}_d \mathbf{n}_d^\dagger] \\ &= \log \det \mathbb{E}[\mathbf{n}_d \mathbf{n}_d^\dagger + \mathbf{h}_1 \mathbf{h}_1^\dagger |x_{r1}|^2 + \mathbf{h}_2 \mathbf{h}_2^\dagger |x_{r2}|^2 + \mathbf{h}_1 \mathbf{h}_2^\dagger x_{r1} x_{r2}^* + \mathbf{h}_2 \mathbf{h}_1^\dagger x_{r1}^* x_{r2}] \\ &= \log \det(\mathbf{I} + \mathbf{h}_1 \mathbf{h}_1^\dagger P_{r1} + \mathbf{h}_2 \mathbf{h}_2^\dagger P_{r2} + \mathbf{h}_1 \mathbf{h}_2^\dagger \rho \sqrt{P_{r1} P_{r2}} + \mathbf{h}_2 \mathbf{h}_1^\dagger \rho^* \sqrt{P_{r1} P_{r2}}) \\ &= \log \det(\mathbf{I} + \mathbf{H} \mathbf{Q}_r \mathbf{H}^\dagger). \end{aligned}$$

where $\rho = \frac{\mathbb{E}[x_{r1} x_{r2}^*]}{\sqrt{P_{r1} P_{r2}}}$ and note that we set $\angle \rho = \angle \mathbf{h}_1^\dagger \mathbf{h}_2$ in (2) and (3). From the third cut, the following upper bounds are derived:

$$\begin{aligned} I(\mathbf{x}_s; y_{r2}) + I(x_{r1}; \mathbf{y}_d | x_{r2}) &\leq \log(1 + \mathbf{g}_2 \mathbf{Q}_s \mathbf{g}_2^\dagger) + \log(1 + (1 - |\rho|^2) P_{r1} \|\mathbf{h}_1\|^2) \\ &\leq \log(1 + P_s \|\mathbf{g}_2\|^2) + \log(1 + (1 - |\rho|^2) P_{r1} \|\mathbf{h}_1\|^2) \end{aligned}$$

where the first inequality follows from the fact that the circularly symmetric complex Gaussian distribution maximized differential entropy and from the following property:

$$\mathbb{E}_{x_{r2}} [\mathbb{E}_{x_{r1}} \{ |x_{r1}|^2 |x_{r2} \} - |\mathbb{E}_{x_{r1}} \{ x_{r1} | x_{r2} \}|^2] \leq (1 - |\rho|^2) P_{r1},$$

and for the second inequality $\mathbf{Q}_s = \frac{P_s}{\|\mathbf{g}_2\|^2} \mathbf{g}_2 \mathbf{g}_2^\dagger$ is used. Similarly, from the fourth cut:

$$\begin{aligned} I(\mathbf{x}_s; y_{r1}) + I(x_{r2}; \mathbf{y}_d | x_{r1}) &\leq \log(1 + \mathbf{g}_1 \mathbf{Q}_s \mathbf{g}_1^\dagger) + \log(1 + (1 - |\rho|^2) P_{r2} \|\mathbf{h}_2\|^2) \\ &\leq \log(1 + P_s \|\mathbf{g}_1\|^2) + \log(1 + (1 - |\rho|^2) P_{r2} \|\mathbf{h}_2\|^2). \end{aligned}$$

By properly combining the above four bounds, we can see that the minimum of the three expressions given in the theorem statement results in a tighter upper bound. \blacksquare

IV. DECODE-AND-FORWARD (DF)

In this section, we describe an achievable rate of the vector Gaussian parallel relay network with a DF strategy. With DF, the signal is delivered from the source to the destination in a two hop transmission. The source has three message sets: $m_1 \in \{1, 2, \dots, 2^{nR_1}\}$ intended for relay 1, $m_2 \in \{1, 2, \dots, 2^{nR_2}\}$ for relay 2, and $m_c \in \{1, 2, \dots, 2^{nR_c}\}$ for both relays. Depending on the messages to transmit, the source node makes the signal \mathbf{x}_s as a function of m_1 , m_2 and m_c . Upon successful decoding of the signals, relays know their private messages and the common message. The relays re-encode the received information to make input signals x_{r1} as a function of m_1 and m_c , and x_{r2} as a function of m_2 and m_c as a block Markov manner. As the end-to-end channel is a cascade of MISO BC and SIMO MAC both with common information, we first investigate optimal signaling in the second hop.

A. The Second Hop: SIMO MAC with Common Information

The capacity region of the discrete memoryless MAC with common information is derived in [14], [16] and that of scalar Gaussian MAC with common information in [17]. The characteristics of the scalar Gaussian MAC were further investigated in [18]. The result can be extended to the Gaussian MAC with multiple antennas at the destination.

Definition 4: A $((2^{nR_1} \times 2^{nR_c}, 2^{nR_2} \times 2^{nR_c}), n)$ code for the SIMO MAC with common information consists of three message sets $m_1 = \{1, 2, \dots, 2^{nR_1}\}$, $m_2 = \{1, 2, \dots, 2^{nR_2}\}$, $m_c = \{1, 2, \dots, 2^{nR_c}\}$, encoding functions at two relays, $f_{r1} : (m_1, m_c) \rightarrow \mathbb{C}^n$ and $f_{r2} : (m_2, m_c) \rightarrow \mathbb{C}^n$, respectively, and a decoding function at the destination $g_d : \mathbb{C}^{M \times n} \rightarrow (\hat{m}_1, \hat{m}_2, \hat{m}_c)$. If the messages (m_1, m_2, m_c) are sent, the conditional probability of error is $\lambda(m_1, m_2, m_c) = \Pr\{g_d(\mathbf{y}_d^n) \neq (m_1, m_2, m_c) | (m_1, m_2, m_c) \text{ sent}\}$. The average probability of error is defined as $P_e^{(n)} = \frac{1}{2^{n(R_1+R_2+R_c)}} \sum_{(m_1, m_2, m_c)} \lambda(m_1, m_2, m_c)$.

Definition 5: If there exists a sequence of $((2^{nR_1} \times 2^{nR_c}, 2^{nR_2} \times 2^{nR_c}), n)$ codes with $P_e^{(n)} \rightarrow 0$, the rate triplet (R_1, R_2, R_c) is said to be achievable.

Definition 6: The capacity region of the SIMO MAC with common information is the closure of the set of all achievable rate triplets.

We will derive an achievable rate region assuming Gaussian input distributions. Each relay's input signal is a superposition of private and common signals:

$$x_{r1}^n(m_1, m_c) = v^n(m_1) + \sqrt{\frac{\alpha P_{r1}}{P_c}} u_1^n(m_c), \quad x_{r2}^n(m_2, m_c) = w^n(m_2) + \sqrt{\frac{\beta P_{r2}}{P_c}} u_2^n(m_c)$$

where $v \sim \mathcal{CN}(0, \bar{\alpha} P_{r1})$, $w \sim \mathcal{CN}(0, \bar{\beta} P_{r2})$, $\bar{\alpha} = 1 - \alpha$, $\bar{\beta} = 1 - \beta$, $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, and u_1 and u_2 are partially correlated random variables in the sense that

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \sim \mathcal{CN}\left(0, P_c \begin{bmatrix} 1 & \gamma \\ \gamma^* & 1 \end{bmatrix}\right).$$

The conditional probability density functions are given by

$$p(x_{r1}|u_1) \sim \mathcal{CN}\left(\sqrt{\frac{\alpha P_{r1}}{P_c}} u_1, \bar{\alpha} P_{r1}\right), \quad p(x_{r2}|u_2) \sim \mathcal{CN}\left(\sqrt{\frac{\beta P_{r2}}{P_c}} u_2, \bar{\beta} P_{r2}\right).$$

Note that the correlation coefficient between the relay input signals is given by $\rho = \frac{\mathbb{E}[x_{r1}x_{r2}^*]}{\sqrt{P_{r1}P_{r2}}} = \gamma\sqrt{\alpha\beta}$ and controllable by power allocation. The received signal at the destination is given by

$$\begin{aligned} \mathbf{y}_d &= \mathbf{h}_1 x_{r1} + \mathbf{h}_2 x_{r2} + \mathbf{n}_d \\ &= \mathbf{h}_1 \left(v + \sqrt{\frac{\alpha P_{r1}}{P_c}} u_1\right) + \mathbf{h}_2 \left(w + \sqrt{\frac{\beta P_{r2}}{P_c}} u_2\right) + \mathbf{n}_d \\ &= \mathbf{h}_1 v + \mathbf{h}_2 w + \left(\mathbf{h}_1 \sqrt{\frac{\alpha P_{r1}}{P_c}} u_1 + \mathbf{h}_2 \sqrt{\frac{\beta P_{r2}}{P_c}} u_2\right) + \mathbf{n}_d. \\ &= \mathbf{h}_1 v + \mathbf{h}_2 w + \left[\mathbf{h}_1 \sqrt{\frac{\alpha P_{r1}}{P_c}} \quad \mathbf{h}_2 \sqrt{\frac{\beta P_{r2}}{P_c}}\right] \mathbf{u} + \mathbf{n}_d. \end{aligned} \tag{4}$$

The rate region of the SIMO Gaussian MAC in terms of mutual information expression can be written as follows

$$\mathcal{R}_{MAC} = \left\{ (R_1, R_2, R_c) \left| \begin{array}{l} R_1 \leq I(x_{r1}; \mathbf{y}_d | x_{r2}, u_1, u_2), \\ R_2 \leq I(x_{r2}; \mathbf{y}_d | x_{r1}, u_1, u_2), \\ R_1 + R_2 \leq I(x_{r1}, x_{r2}; \mathbf{y}_d | u_1, u_2), \\ R_1 + R_2 + R_c \leq I(x_{r1}, x_{r2}; \mathbf{y}_d) \end{array} \right. \right\}$$

for some distribution $p(u_1, u_2)p(x_{r1}|u_1)p(x_{r2}|u_2)$ where u_1 and u_2 are auxiliary random variables that represent common information.

Theorem 2: The following rate region is achievable for the SIMO MAC with common information:

$$\mathcal{R}_{MAC}(P_{r1}, P_{r2}) = \bigcup_{\alpha, \beta, \mathbb{E}[|x_i|^2] \leq P_{ri}, i=1,2} \mathcal{R}(\alpha, \beta, \gamma) \tag{5}$$

where $\mathcal{R}(\alpha, \beta, \gamma)$ is the rate region for given α, β and γ , which can be expressed as

$$\mathcal{R}(\alpha, \beta, \gamma) = \left\{ (R_1, R_2, R_c) \left| \begin{array}{l} R_1 \leq \log(1 + \overline{\alpha}P_{r1}\|\mathbf{h}_1\|^2), \\ R_2 \leq \log(1 + \overline{\beta}P_{r2}\|\mathbf{h}_2\|^2), \\ R_1 + R_2 \leq \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r^p\mathbf{H}^\dagger), \\ R_1 + R_2 + R_c \leq \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) \end{array} \right. \right\} \quad (6)$$

where

$$\begin{aligned} \mathbf{Q}_r^p &= \mathbb{E}_u [\mathbb{E}_{\mathbf{x}_r} [\mathbf{x}_r \mathbf{x}_r^\dagger | u]] = \begin{bmatrix} \overline{\alpha}P_{r1} & 0 \\ 0 & \overline{\beta}P_{r2} \end{bmatrix} \\ \mathbf{Q}_r &= \mathbb{E}_{\mathbf{x}_r} [\mathbf{x}_r \mathbf{x}_r^\dagger] = \begin{bmatrix} P_{r1} & \gamma\sqrt{\alpha\beta P_{r1}P_{r2}} \\ \gamma^*\sqrt{\alpha\beta P_{r1}P_{r2}} & P_{r2} \end{bmatrix}. \end{aligned}$$

Proof: It is straightforward to show the theorem result by evaluating the mutual information expressions assuming circularly symmetric complex Gaussian input distributions. ■

Next, we are interested in how to maximize the sum-rate and get the following result for the optimal correlation that maximizes $\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger)$.

Lemma 1 (Optimal correlation): For any N , the sum rate of the SIMO MAC with common information is maximized by $\angle\rho = \angle\gamma = \angle\mathbf{h}_1^\dagger\mathbf{h}_2$ and $|\rho| = |\gamma|\sqrt{\alpha\beta} = \min\left(\frac{|\mathbf{h}_1^\dagger\mathbf{h}_2|}{\sqrt{P_{r1}P_{r2}}\det(\mathbf{H}^\dagger\mathbf{H})}, 1\right)$, and the resulting maximum sum-rate is given by

$$\begin{aligned} R_{sum, MAC}^{max} &= \max_{\rho} \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) \\ &= \begin{cases} \log\left(1 + P_{r1}\|\mathbf{h}_1\|^2 + P_{r2}\|\mathbf{h}_2\|^2 + P_{r1}P_{r2}\det(\mathbf{H}^\dagger\mathbf{H}) + \frac{|\mathbf{h}_1^\dagger\mathbf{h}_2|^2}{\det(\mathbf{H}^\dagger\mathbf{H})}\right), & |\rho|^{opt} < 1 \\ \log\left(1 + P_{r1}\|\mathbf{h}_1\|^2 + P_{r2}\|\mathbf{h}_2\|^2 + 2\sqrt{P_{r1}P_{r2}}|\mathbf{h}_1^\dagger\mathbf{h}_2|\right), & |\rho|^{opt} = 1. \end{cases} \end{aligned}$$

Proof: See Appendix. ■

Remark 1: By defining the angle between channel vectors $\varphi_h = \arccos \frac{|\mathbf{h}_1^\dagger\mathbf{h}_2|}{\|\mathbf{h}_1\|\|\mathbf{h}_2\|}$, $\varphi_h \in [0, \pi/2]$, and the geometric mean $\text{SNR}_{geo} = \sqrt{P_{r1}P_{r2}}\|\mathbf{h}_1\|\|\mathbf{h}_2\|$, the optimal correlation can be also expressed as

$$|\rho|^{opt} = \min\left(\frac{\cos\varphi_h}{\text{SNR}_{geo}\sin^2\varphi_h}, 1\right).$$

It is a monotonically decreasing function of φ_h and inversely proportional to SNR_{geo} . In Fig. 3, $|\rho|^{opt}$ is drawn for different values of φ_h and SNR_{geo} .

For a fixed SNR_{geo} or for a fixed φ_h , there exists a threshold above which $|\rho|^{opt} < 1$ and below which $|\rho|^{opt} = 1$. The threshold is the solution of $\cos\varphi_h = \text{SNR}_{geo}\sin^2\varphi_h$.

Remark 2: If \mathbf{h}_1 and \mathbf{h}_2 are orthogonal, then $|\rho|^{opt} = 0$. If channel vectors are orthogonal, the differential entropy of the received signal vector \mathbf{y}_d is maximized when x_{r1} and x_{r2} are uncorrelated. The resulting sum-rate is given by

$$\max_{\rho} \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) = \log(1 + P_{r1}\|\mathbf{h}_1\|^2) + \log(1 + P_{r2}\|\mathbf{h}_2\|^2).$$

If \mathbf{h}_1 and \mathbf{h}_2 are parallel, then $|\rho|^{opt} = 1$. If channel vectors are parallel, the differential entropy of the received signal vector \mathbf{y}_d is maximized when x_{r1} and x_{r2} are fully correlated. The resulting sum-rate is given by

$$\max_{\rho} \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) = \log(1 + P_{r1}\|\mathbf{h}_1\|^2 + P_{r2}\|\mathbf{h}_2\|^2 + 2\sqrt{P_{r1}P_{r2}}|\mathbf{h}_1^\dagger\mathbf{h}_2|).$$

Remark 3: For a fixed \mathbf{H} , if either P_{r1} or P_{r2} are sufficiently small so that $|\rho|^{opt} = 1$, the signaling is optimal when the relay signals are perfectly correlated. If $P_{r1}, P_{r2} > 0$ and either P_{r1} or P_{r2} are sufficiently large so that $|\rho|^{opt} = \frac{|\mathbf{h}_1^\dagger\mathbf{h}_2|}{\sqrt{P_{r1}P_{r2}}\det(\mathbf{H}^\dagger\mathbf{H})} < 1$, then $|\rho|^{opt} \rightarrow 0$ as either P_{r1} or $P_{r2} \rightarrow \infty$. If power is abundant, very small fraction of relay power needs to be allocated to common signals to satisfy optimality condition.

Remark 4: For $N = 1$, the sum rate is maximized by $|\rho| = 1$, i.e., $\alpha = \beta = |\gamma| = 1$ and $\angle\rho = \angle(h_1^*h_2)$ where $\mathbf{H} = [h_1 \ h_2]$.

By combining the optimal correlation condition with the achievable region expression in (6), we get the following result.

Theorem 3 (Maximum sum-rate subregion): In the three-dimensional achievable region of a SIMO MAC given by (6), the maximum sum-rate subregion is a surface whose boundary is characterized by

$$\mathcal{R}_{sub}(|\rho|^{opt}) = \bigcup_{|\rho|^{opt} \leq \sqrt{\alpha} \leq 1} \mathcal{R}(\alpha) \quad (7)$$

where

$$\mathcal{R}(\alpha) = \left\{ (R_1, R_2, R_c) \left| \begin{array}{l} R_1 \leq \log(1 + \bar{\alpha}P_{r1}\|\mathbf{h}_1\|^2) \\ R_2 \leq \log(1 + (1 - (|\rho|^{opt}/\sqrt{\alpha})^2)P_{r2}\|\mathbf{h}_2\|^2) \text{ for } |\rho|^{opt} > 0 \\ R_2 \leq \log(1 + P_{r2}\|\mathbf{h}_2\|^2) \text{ for } |\rho|^{opt} = 0 \\ R_1 + R_2 \leq \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r^p\mathbf{H}^\dagger) \\ R_1 + R_2 + R_c = \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) \end{array} \right. \right\} \quad (8)$$

where

$$\mathbf{Q}_r^p = \begin{bmatrix} (1 - \alpha)P_{r1} & 0 \\ 0 & (1 - (|\rho|^{opt}/\sqrt{\alpha})^2)P_{r2} \end{bmatrix} \text{ for } |\rho|^{opt} > 0,$$

$$\mathbf{Q}_r^p = \begin{bmatrix} (1 - \alpha)P_{r1} & 0 \\ 0 & P_{r2} \end{bmatrix} \text{ for } |\rho|^{opt} = 0,$$

$$\mathbf{Q}_r = \begin{bmatrix} P_{r1} & |\rho|^{opt}e^{j\angle\mathbf{h}_1^\dagger\mathbf{h}_2}\sqrt{P_{r1}P_{r2}} \\ |\rho|^{opt}e^{-j\angle\mathbf{h}_1^\dagger\mathbf{h}_2}\sqrt{P_{r1}P_{r2}} & P_{r2} \end{bmatrix}.$$

Proof: To satisfy the optimal correlation, $0 \leq |\rho|^{opt} = |\gamma|\sqrt{\alpha\beta} \leq 1$, $|\gamma|$, $\sqrt{\alpha}$ and $\sqrt{\beta}$ should be in the range $[\|\rho\|^{opt}, 1]$. In (6), by setting $|\gamma| = 1$, and $\beta = (|\rho|^{opt}/\sqrt{\alpha})^2$ for $|\rho|^{opt} > 0$ or $\beta = 0$ for $|\rho|^{opt} = 0$, and by taking union over $|\rho|^{opt} \leq \sqrt{\alpha} \leq 1$, we characterize the boundary of the maximum sum-rate surface. \blacksquare

Example 1 (Close-to-parallel channel vectors): If $|\rho|^{opt} = 1$, it is required for relays to set $\alpha = \beta = |\gamma| = 1$. Substituting this condition into (6), we get the expression for the maximum sum-rate subregion which is a single point given by

$$\left\{ (R_1, R_2, R_c) \left| \begin{array}{l} R_1 = 0, R_2 = 0, \\ R_c = \log(1 + P_{r1} \|\mathbf{h}_1\|^2 + P_{r2} \|\mathbf{h}_2\|^2 + 2\sqrt{P_{r1}P_{r2}}|\mathbf{h}_1^\dagger\mathbf{h}_2|) \end{array} \right. \right\}.$$

The example achievable rate region is drawn in Fig. 4 (a) and (b), and the maximum sum-rate point $(0, 0, R_{sum,MAC}^{max})$ is on the R_c axis.

Example 2 (Orthogonal channel vectors): If $|\rho|^{opt} = 0$, at least one of α , β and $|\gamma|$ must be zero. We get the expression for the maximum sum-rate subregion given by

$$\left\{ (R_1, R_2, R_c) \left| \begin{array}{l} R_1 \leq \log(1 + P_{r1} \|\mathbf{h}_1\|^2) \\ R_2 \leq \log(1 + P_{r2} \|\mathbf{h}_2\|^2) \\ R_1 + R_2 + R_c \leq \sum_{i=1,2} \log(1 + P_{ri} \|\mathbf{h}_i\|^2) \end{array} \right. \right\}.$$

As it is shown in Fig. 4 (d), the maximum sum-rate surface is the square connecting four points: $(R_1^{max}, R_2^{max}, 0)$, $(R_1^{max}, 0, R_{sum,MAC}^{max} - R_1^{max})$, $(0, R_2^{max}, R_{sum,MAC}^{max} - R_2^{max})$, and $(0, 0, R_{sum,MAC}^{max})$, where R_i^{max} denotes the maximum rate that can be achieved by each message set. In the subregion, for any fixed R_1 and R_2 , we can find R_c such that $R_{sum,MAC}^{max} = R_1 + R_2 + R_c$.

B. Impact of Common Information Signaling

Let us discuss how much benefit we can get by having the common information for $N \geq 2$ case. If the source transmits only the private signal to relays, i.e. $R_c = 0$, the best strategy relays can do is to have diagonal covariance matrix with individual peak power $\mathbf{Q}_r^{diag} = diag[P_{r1}, P_{r2}]$. Let \mathbf{Q}_r^{opt} denote optimal covariance matrix with sum rate maximizing magnitude and phase angle of ρ derived above. Then, we get the following result.

Lemma 2 (Benefit of correlation): $\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r^{diag}\mathbf{H}^\dagger) \leq \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r^{opt}\mathbf{H}^\dagger)$ with equality if and only if $\mathbf{h}_1^\dagger \mathbf{h}_2 = 0$. The increase in SNR by having \mathbf{Q}_r^{opt} is given by

$$\Delta \text{SNR} = \begin{cases} \frac{|\mathbf{h}_1^\dagger \mathbf{h}_2|^2}{\det(\mathbf{H}^\dagger \mathbf{H})}, & |\rho|^{opt} < 1 \\ 2\sqrt{P_{r1}P_{r2}}|\mathbf{h}_1^\dagger \mathbf{h}_2|, & |\rho|^{opt} = 1 \end{cases}.$$

Proof: : It is sufficient to show that the sum-rate is a quadratic and concave function of $|\rho|$, and is monotonically increasing for $0 \leq |\rho| \leq |\rho|^{opt}$. The function has its minimum at $|\rho| = 0$ since $|\rho|$ is non-negative. When the channel vectors are orthogonal, the suboptimality vanishes since $|\rho|^{opt} = 0$. The SNR increase can be directly calculated using the result in Lemma 1. ■

When the channel column vectors \mathbf{h}_1 and \mathbf{h}_2 are close to orthogonal, \mathbf{Q}_r^{diag} is almost as good as \mathbf{Q}_r^{opt} . However, when \mathbf{h}_1 and \mathbf{h}_2 are close to parallel, the sum rate by having \mathbf{Q}_r^{opt} at relays shows considerable increase from that by having \mathbf{Q}_r^{diag} . The gain coming from optimal correlation becomes very large at low SNR. Fig. 5 shows the examples.

With common information coming from the source, we can introduce correlation between relays, and they act as if they are in cooperation. The resulting SIMO MAC behaves like a point-to-point MIMO channel with per-antenna power constraint.

Here, we can see that there is a minimum required R_c that needs to be transmitted from the source to relays for achieving maximum sum-rate in the second hop.

Theorem 4 (Threshold of R_c): In the SIMO MAC with common information, the threshold of R_c above which a maximum sum-rate point can exist is characterized by

$$R_c^{th} = \log \frac{\det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r^{opt}\mathbf{H}^\dagger)}{\det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r^k\mathbf{H}^\dagger)}$$

where $\mathbf{Q}_r^k = diag[(1 - k|\rho|^{opt})P_{r1}, (1 - k^{-1}|\rho|^{opt})P_{r2}]$ and $k = \sqrt{\frac{P_{r2}\|\mathbf{h}_2\|^2 + P_{r1}P_{r2}\det(\mathbf{H}^\dagger \mathbf{H})}{P_{r1}\|\mathbf{h}_1\|^2 + P_{r1}P_{r2}\det(\mathbf{H}^\dagger \mathbf{H})}}$.

Proof: In the maximum sum-rate subregion in (8), after evaluating $\det(\cdot)$ operation, we get the expression for $R_1 + R_2$ given by

$$R_1 + R_2 \leq \log \left(1 + \overline{\alpha} P_{r1} \|\mathbf{h}_1\|^2 + \left(1 - \frac{|\rho|^2}{\alpha} \right) P_{r2} \|\mathbf{h}_2\|^2 + \left(1 - \alpha + |\rho|^2 - \frac{|\rho|^2}{\alpha} \right) P_{r1} P_{r2} \det(\mathbf{H}^\dagger \mathbf{H}) \right).$$

It is straightforward to check that the maximum of $R_1 + R_2$ is given by

$$\max_{\alpha} (R_1 + R_2) = \log \det(\mathbf{I} + \mathbf{H} \mathbf{Q}_r^k \mathbf{H}^\dagger)$$

with \mathbf{Q}_r^k given in the theorem statement. Finally, $R_c^{th} = R_{sum,MAC}^{max} - \max_{\alpha} (R_1 + R_2)$ results in the minimum possible value of R_c while staying in the maximum sum-rate subregion. ■

C. SISO MAC versus SIMO MAC

Optimal signaling at DF relays depends on the channel condition. For a SISO MAC with a single antenna at the destination ($N = 1$), the sum-rate is maximized by $\alpha = \beta = |\gamma| = 1$ so that $R_1 = R_2 = 0$ regardless of the channel and power constraints. This is the case when all the power is allocated to the common signal at both relays and they add up coherently at the destination. It is desired for the source to transmit as much common information to relays as possible, and this strategy maximizes the source-to-destination sum-rate. Even when $N \geq 2$, if \mathbf{h}_1 and \mathbf{h}_2 are close to parallel in the sense that $\det(\mathbf{H}^\dagger \mathbf{H})$ is small and $|\rho|^{opt} = 1$, fully correlated relay signals are still optimal.

In contrast, with multiple antennas at the destination ($N \geq 2$) and $|\rho|^{opt} < 1$, sum-rate maximizing α and β depend on both the channel matrix \mathbf{H} and relay power constraints. Any combination of power allocation factors at relays such that $|\gamma| \sqrt{\alpha \beta} = \frac{|\mathbf{h}_1^\dagger \mathbf{h}_2|}{\sqrt{P_{r1} P_{r2}} \det(\mathbf{H}^\dagger \mathbf{H})} < 1$ together with optimal rotation angle $\angle \gamma = \angle \mathbf{h}_1^\dagger \mathbf{h}_2$ can maximize the sum rate of a SIMO MAC. In this case, the source needs to transmit just the right amount of common information so that signals at two relays are optimally correlated and the source-to-destination sum-rate is maximized.

D. The First Hop: MISO BC with Common Information

In [1], the first hop is a degraded scalar BC where one relay with higher SNR can decode both its intended signal and the signal for the other relay by doing superposition coding and

successive interference cancellation. In this case, correlation between relay input signals are naturally introduced. In contrast, our first hop is a non-degraded vector broadcast channel that makes it possible to send private signals, each decodable by one of the relays, as well as a common signal decodable by both relays. For this class of channels, the three dimensional capacity region is not known, but a good achievable region combining dirty paper coding (DPC) [13] and superposition was studied in [10]–[12] which we also use here.

In this scheme, the transmitting signal \mathbf{x}_s is a superposition of three independent signals \mathbf{x}_1 , \mathbf{x}_2 and \mathbf{x}_c , i.e., $\mathbf{x}_s = \mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_c$, where $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_c \in \mathbb{C}^{N \times 1}$ denote the signals intended for relay 1, for relay 2 and for both relays, i.e., the common message, respectively. We assume Gaussian signaling for all signals. Input covariance matrix is $\mathbf{Q}_s = \mathbf{Q}_1 + \mathbf{Q}_2 + \mathbf{Q}_c$, where $\mathbf{Q}_j = \mathbb{E}[\mathbf{x}_j \mathbf{x}_j^\dagger]$, $j \in \{1, 2, c\}$.

Common information is decoded at both relays before decoding private messages. Private messages are encoded using dirty paper coding, i.e., the private message for relay 1 is first encoded as \mathbf{x}_1 , and the private message for relay 2 is then encoded as \mathbf{x}_2 using \mathbf{x}_1 as side information so that \mathbf{x}_2 can be decoded at relay 2 without interference from \mathbf{x}_1 . The encoding order can be reversed. With this scheme, an achievable rate region is given by

$$\mathcal{R}_{BC}(P_s) = Co \left(\bigcup_{\pi, \mathbf{Q}_s: \text{tr}(\mathbf{Q}_s) \leq P_s} \mathcal{R}(\pi, \mathbf{Q}_s) \right) \quad (9)$$

where $\mathcal{R}(\pi, \mathbf{Q}_s)$ is the achievable region for a given encoding order $\pi \in \{\pi_{12}, \pi_{21}\}$ and input covariance matrix \mathbf{Q}_s , where $Co(\cdot)$ is the convex hull operator. If \mathbf{x}_2 is encoded first, for example, then we have

$$\mathcal{R}(\pi_{12}, \mathbf{Q}_s) = \left\{ (R_1, R_2, R_c) \left| \begin{array}{l} R_1 \leq \log \left(1 + \mathbf{g}_1 \mathbf{Q}_1 \mathbf{g}_1^\dagger \right), \\ R_2 \leq \log \left(1 + \frac{\mathbf{g}_2 \mathbf{Q}_2 \mathbf{g}_2^\dagger}{1 + \mathbf{g}_2 \mathbf{Q}_1 \mathbf{g}_2^\dagger} \right), \\ R_c \leq \min_{i \in \{1, 2\}} \log \left(1 + \frac{\mathbf{g}_i \mathbf{Q}_c \mathbf{g}_i^\dagger}{1 + \mathbf{g}_i (\mathbf{Q}_1 + \mathbf{Q}_2) \mathbf{g}_i^\dagger} \right) \end{array} \right. \right\}. \quad (10)$$

Fig. 6. (a) depicts an example of an achievable region where two row vectors of \mathbf{G} are parallel and linearly dependent so that \mathbf{G} is ill-conditioned and rank-deficient. Note that in the figures, $\varphi_g = \arccos \frac{|\mathbf{g}_1 \mathbf{g}_2^\dagger|}{\|\mathbf{g}_1\| \|\mathbf{g}_2\|}$. One can see that the maximum sum-rate surface is a plane connecting three points: $(R_{1,BC}^{max}, 0, 0)$, $(0, R_{2,BC}^{max}, 0)$, and $(0, 0, R_{c,BC}^{max})$ where $R_{j,BC}^{max}$ denotes the maximum rate achieved by allocating all power to \mathbf{x}_j so that $\mathbf{Q}_j = \mathbf{Q}_s$:

$$R_{i,BC}^{max} = \max_{\mathbf{Q}_s} \log \left(1 + \mathbf{g}_i \mathbf{Q}_s \mathbf{g}_i^\dagger \right) = \log \left(1 + P_s \|\mathbf{g}_i\|^2 \right)$$

$$R_{c,BC}^{max} = \max_{\mathbf{Q}_s} \min_i \log(1 + \mathbf{g}_i \mathbf{Q}_s \mathbf{g}_i^\dagger) = \min_i \log(1 + P_s \|\mathbf{g}_i\|^2)$$

where $i \in \{1, 2\}$. In this specific example, the channel is symmetric in the sense that $\|\mathbf{g}_1\| = \|\mathbf{g}_2\|$. In fact, this is the only case where having common information does not incur sum-rate penalty.

As the opposite extreme, Fig. 6. (b) depicts an example of an achievable region where two row vectors of \mathbf{G} are orthogonal and linearly independent so that \mathbf{G} is well-conditioned and full-rank. In the symmetric example, the point that achieves the maximum sum-rate

$$R_{sum,BC}^{max} = \log(1 + \frac{P_s}{2} \|\mathbf{g}_1\|) + \log(1 + \frac{P_s}{2} \|\mathbf{g}_2\|)$$

is on the line $R_1 = R_2$ and $R_c = 0$. The point that has the minimum sum-rate on the boundary is on the R_c axis, i.e., $R_1 = R_2 = 0$ and

$$R_{c,BC}^{max} = \max_{\mathbf{Q}_s} \min_i \log(1 + \mathbf{g}_i \mathbf{Q}_s \mathbf{g}_i^\dagger) = \log(1 + \frac{P_s}{2} \|\mathbf{g}_1\|) = \log(1 + \frac{P_s}{2} \|\mathbf{g}_2\|)$$

where we can see the sum-rate penalty due to beamforming inefficiency.

Note that the maximum sum-rate points of a MISO BC are always on the $R_1 - R_2$ plane. In fact, they correspond to the dominant face of the two user BC achievable region by DPC without common information.

E. Achievable Rate by DF

For the GPRN drawn in Fig. 1, a triplet (R_1, R_2, R_c) is said to be achievable by DF if it belongs to the intersection of the rate regions of the first hop MISO BC and the second hop SIMO MAC. In this context, the maximum rate by DF can be defined by

$$R_{DF}^{max} = \max_{(R_1, R_2, R_c) \in \mathcal{R}_{DF}} R_1 + R_2 + R_c \quad (11)$$

where $\mathcal{R}_{DF} = \mathcal{R}_{BC}(P_s) \cap \mathcal{R}_{MAC}(P_{r1}, P_{r2})$. Fig. 7. shows examples of the rate regions of MISO BC and SIMO MAC, the intersection of which is the achievable rate region by DF.

If the source-to-relay link SNR is high enough, the second hop becomes the bottleneck and determines the source-to-destination sum-rate.

Theorem 5 (Optimality condition of DF): If there is a rate triple $(R_1, R_2, R_c) \in \mathcal{R}_{sub}(|\rho|^{opt})$ which is included in the MISO BC region $\mathcal{R}_{BC}(P_s)$, then $R_{DF}^{max} = R_{sum,MAC}^{max}$ meets the upper bound and determines the capacity of the vector Gaussian parallel relay network.

Proof: Let us first consider $|\rho|^{opt} = 1$ case where there is a single maximum sum-rate point in the SIMO MAC region. At $|\rho| = 1$, the term $\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger)$ in (2) and (3) is maximized and is smaller than (1) since the following relationships are hold:

$$R_{sum,MAC}^{max} \leq R_{c,BC}^{max} \leq R_{i,BC}^{max} (i \in \{1, 2\}) \leq R_{sum,BC}^{max} \leq R_{BC}^{upper}.$$

We can achieve the tightest upper bound $R_{sum,MAC}^{max}$ by allocating all relay power to the common signal: $\alpha = \beta = |\gamma| = 1$.

Now we consider $|\rho|^{opt} < 1$ case where there exist more than one maximum sum-rate point in the SIMO MAC region. It is tedious but easy to verify that, at $|\rho| = |\rho|^{opt}$,

$$R_{sum,MAC}^{max} \leq R_{sum,BC}^{max} \leq R_{BC}^{upper}$$

$$2 \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) \leq \sum_{i=1,2} \log(1 + P_s \|\mathbf{g}_i\|^2) + \sum_{i=1,2} \log(1 + (1 - |\rho|^2)P_{ri} \|\mathbf{h}_i\|^2).$$

We can achieve the tightest upper bound $R_{sum,MAC}^{max}$ by optimal power allocation at relays such that $|\rho|^{opt} = |\gamma| \sqrt{\alpha\beta}$. ■

With DF, the first and second hops are completely separated in the sense that after finishing the first stage of transmission, relays start a new stage of transmission by encoding the received information again. Here, the right approach is first to figure out what is optimal in the second hop, and then to check if the optimal operating point, i.e., one of the maximum sum-rate points of the SIMO MAC is supportable by the first hop. If the optimal point is achievable, the network nodes would start communication by setting parameters to satisfy optimality conditions. If none of the maximum sum-rate points of the SIMO MAC is achievable, then the nodes would try to find the operating point as close to the optimal as possible.

What if the relay-to-destination link SNRs are high enough so that the first hop is the bottleneck? If one of the maximum sum-rate points of the MISO BC on the $R_1 - R_2$ plane is included in \mathcal{R}_{MAC} , then $R_{DF}^{max} = R_{sum,BC}^{max}$. In this case, the broadcast cut-set upper bound in (1) is tighter than the others. It turns out that the upper bound and the achievable rate meet in some special cases where the first hop row vectors are orthogonal as will be seen in numerical results in Section VII. However, they do not meet in general, which implies suboptimality of DF in the case. If the first hop is the bottleneck, full decoding at relays gives too much restriction, and AF and CF schemes can do better.

F. Symmetric Channels

In this subsection, we narrow down our attention to the symmetric channels: $\|\mathbf{g}_1\| = \|\mathbf{g}_2\|$, $\|\mathbf{h}_1\| = \|\mathbf{h}_2\|$, $P_{r1} = P_{r2}$, and $R_1 = R_2$. We consider four examples illustrated in Fig. 8. where $R_p = R_1 + R_2$.

In Fig. 8. (a), the straight line \overline{AB} and the curved line \overline{BC} constitute the surface of the MAC achievable rate region. The curved line \overline{EG} is the surface of the BC achievable rate region. Here, we have the source power constraint P_s so large that the upper bound (1) and the second terms of \min in (2) and (3) are loose. In this case, $\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger)$ is the active upper bound, and the sum-rate constraint $R_1 + R_2 + R_c$ is the straight line that goes through the points A , B , and D . The straight line \overline{AB} is the maximum sum-rate subregion of the MAC, and F is the crossing point of the BC and MAC surfaces. All the points on the line \overline{FB} achieve the capacity of the vector Gaussian parallel relay network as they are in the BC and MAC achievable rate regions and meet the sum-rate upper bound.

Other things being equal, the BC rate region with a smaller power constraint is drawn in Fig. 8. (b). In this example, the crossing point F coincides with the point B . Thus, there exists a single capacity achieving point at $B = F$. In Fig. 8. (c), the BC rate region gets even smaller, and the DF maximum sum-rate point F does not meet the sum-rate upper bound. Thus, in this case, DF does not achieve the capacity of the vector Gaussian parallel relay network. Finally, Fig. 8. (d) illustrates the case where the source power constraint P_s is so small that $\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger)$ in the upper bound expression is not active anymore.

V. AMPLIFY-AND-FORWARD (AF)

We have seen that if the relay-to-destination link SNR is high enough, the first hop is the bottleneck and DF does not satisfy the optimality condition that requires at least one maximum sum-rate point of SIMO MAC should be inside MISO BC achievable region. Instead of requiring signals with low SNR to be decoded at relays, it would be better for relays to just forward their received signals to the destination so that the benefit of high SNR in the second hop is maximally utilized.

A. Achievable Rate of AF

The received signal at relays can be expressed in vector notation by

$$\mathbf{y}_r = \mathbf{G}\mathbf{x}_s + \mathbf{n}_r.$$

With AF, the relays just amplify their received signals before forwarding them to the destination, and the transmit signal vector of the relays is given by

$$\mathbf{x}_r = \mathbf{A}\mathbf{y}_r$$

where $\mathbf{A} = \text{diag}[ae^{j\phi}, b]$. The amplification factors should be in the range

$$0 \leq a \leq a^{\text{peak}} = \sqrt{\frac{P_{r1}}{1 + \mathbf{g}_1 \mathbf{Q}_s \mathbf{g}_1^\dagger}}, \quad 0 \leq b \leq b^{\text{peak}} = \sqrt{\frac{P_{r2}}{1 + \mathbf{g}_2 \mathbf{Q}_s \mathbf{g}_2^\dagger}}$$

because of the power constraints at the relays. The received signal at the destination is given by

$$\begin{aligned} \mathbf{y}_d &= \mathbf{H}\mathbf{x}_r + \mathbf{n}_d = \mathbf{H}\mathbf{A}\mathbf{y}_r + \mathbf{n}_d \\ &= \mathbf{H}\mathbf{A}\mathbf{G}\mathbf{x}_s + \mathbf{H}\mathbf{A}\mathbf{n}_r + \mathbf{n}_d \\ &= \mathbf{H}\mathbf{A}\mathbf{G}\mathbf{x}_s + \mathbf{n}_e \\ &= (ae^{j\phi}\mathbf{h}_1\mathbf{g}_1 + b\mathbf{h}_2\mathbf{g}_2)\mathbf{x}_s + (ae^{j\phi}n_{r1}\mathbf{h}_1 + bn_{r2}\mathbf{h}_2 + \mathbf{n}_d) \end{aligned}$$

where \mathbf{n}_e denotes the total effective noise. Since the noises added at relay receivers are also amplified and forwarded through the channel, the noise vector \mathbf{n}_e is spatially non-white in general. Noise covariance matrix is symmetric and can be decomposed as

$$\mathbf{K} = \mathbb{E}[\mathbf{n}_e \mathbf{n}_e^\dagger] = \mathbf{I} + \mathbf{H}\mathbf{A}\mathbf{A}^\dagger\mathbf{H}^\dagger = \mathbf{U}^\dagger \boldsymbol{\Lambda} \mathbf{U}$$

where $\boldsymbol{\Lambda}$ is a diagonal matrix with eigenvalues of \mathbf{K} as its diagonal elements, and \mathbf{U} is a unitary matrix. Then, the channel can be transformed into an equivalent white noise channel given by

$$\boldsymbol{\Lambda}^{-1/2} \mathbf{U} \mathbf{y}_d = \boldsymbol{\Lambda}^{-1/2} \mathbf{U} \mathbf{H} \mathbf{A} \mathbf{G} \mathbf{x}_s + \boldsymbol{\Lambda}^{-1/2} \mathbf{U} \mathbf{n}_e = \mathbf{F} \mathbf{x}_s + \mathbf{n}_w$$

where \mathbf{F} and \mathbf{n}_w denote the effective source-to-destination channel matrix and the white noise vector, respectively. For a fixed \mathbf{A} , this is a point-to-point MIMO channel whose maximum rank is 2 limited by the number of relays. If all the input signal is Gaussian, we get the expression

for an achievable rate by AF given by

$$\begin{aligned}
R_{AF} &\leq I(\mathbf{x}_s; \mathbf{y}_d) \\
&= \log \frac{\det(\mathbf{I} + \mathbf{H}\mathbf{A}\mathbf{A}^\dagger\mathbf{H}^\dagger + \mathbf{H}\mathbf{A}\mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger\mathbf{A}^\dagger\mathbf{H}^\dagger)}{\det(\mathbf{I} + \mathbf{H}\mathbf{A}\mathbf{A}^\dagger\mathbf{H}^\dagger)} \\
&= \log \det(\mathbf{I} + \mathbf{\Lambda}^{-1/2}\mathbf{U}\mathbf{H}\mathbf{A}\mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger\mathbf{A}^\dagger\mathbf{H}^\dagger\mathbf{U}^\dagger\mathbf{\Lambda}^{-1/2}) \\
&= \log \det(\mathbf{I} + \mathbf{F}\mathbf{Q}_s\mathbf{F}^\dagger).
\end{aligned} \tag{12}$$

In order to get the maximum achievable rate, the source signal covariance matrix \mathbf{Q}_s and the relay amplification matrix \mathbf{A} need to be jointly optimized.

Theorem 6 (Asymptotic Optimality of AF): If \mathbf{H} is full rank, AF is asymptotically optimal in the high relay power limit in the sense that

$$\lim_{P_{r1}=P_{r2} \rightarrow \infty} (\max_{\mathbf{Q}_s} \log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger) - \max_{\mathbf{Q}_s, \mathbf{A}} R_{AF}) = 0.$$

Proof: By rearranging (12), we get

$$\begin{aligned}
R_{AF} &= \log \det(\mathbf{I} + \mathbf{H}\mathbf{A}\mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger\mathbf{A}^\dagger\mathbf{H}^\dagger (\mathbf{I} + \mathbf{H}\mathbf{A}\mathbf{A}^\dagger\mathbf{H}^\dagger)^{-1}) \\
&= \log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger\mathbf{A}^\dagger\mathbf{H}^\dagger (\mathbf{I} + \mathbf{H}\mathbf{A}\mathbf{A}^\dagger\mathbf{H}^\dagger)^{-1} \mathbf{H}\mathbf{A}) \\
&= \log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger (\mathbf{I} - (\mathbf{I} + \mathbf{A}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{A})^{-1}))
\end{aligned}$$

where we use matrix inversion lemma. We first set a and b to peak values under relay power constraints, assume ϕ is optimally chosen, and let R_{AF}^{peak} denote the corresponding rate. By showing that if $\det(\mathbf{H}^\dagger\mathbf{H}) > 0$, $(\mathbf{I} + \mathbf{A}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{A})^{-1} \rightarrow \mathbf{0}$ as $a \rightarrow \infty$ and $b \rightarrow \infty$ where

$$(\mathbf{I} + \mathbf{A}^\dagger\mathbf{H}^\dagger\mathbf{H}\mathbf{A})^{-1} = \frac{\begin{bmatrix} 1 + b^2 \|\mathbf{h}_2\|^2 - abe^{-j\phi} \mathbf{h}_1^\dagger \mathbf{h}_2 \\ -abe^{j\phi} \mathbf{h}_2^\dagger \mathbf{h}_1 & 1 + a^2 \|\mathbf{h}_1\|^2 \end{bmatrix}}{1 + a^2 \|\mathbf{h}_1\|^2 + b^2 \|\mathbf{h}_2\|^2 + a^2 b^2 \det(\mathbf{H}^\dagger\mathbf{H})},$$

we get the following result

$$\lim_{P_{r1}=P_{r2} \rightarrow \infty} (\log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger) - R_{AF}^{peak}) = 0$$

Then, the theorem statement naturally follows since by definition, for a fixed \mathbf{Q}_s ,

$$R_{AF}^{peak} \leq \max_{\mathbf{A}} R_{AF}.$$

■

B. Iterative Optimization Algorithm for AF

We can maximize R_{AF} by an iterative algorithm as follows. For optimizing \mathbf{Q}_s , we apply singular value decomposition (SVD) to \mathbf{F} and waterfilling over two parallel scalar channels with non-zero singular values [19]. First, we define the covariance matrix of \mathbf{Gx}_s and the rate function of (a, b) given by

$$\mathbf{Q}^e = \mathbb{E}[\mathbf{Gx}_s \mathbf{x}_s^\dagger \mathbf{G}^\dagger] = \mathbf{GQ}_s \mathbf{G}^\dagger = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix},$$

$$R(a, b) =$$

$$\log \left(1 + \frac{a^2 b^2 \det(\mathbf{H}^\dagger \mathbf{H})(\text{tr}(\mathbf{Q}^e) + \det(\mathbf{Q}^e)) + a^2 \|\mathbf{h}_1\|^2 q_{11} + b^2 \|\mathbf{h}_2\|^2 q_{22} + 2ae^{-j\phi} b \mathbf{h}_1^\dagger \mathbf{h}_2 q_{21}}{1 + a^2 \|\mathbf{h}_1\|^2 + b^2 \|\mathbf{h}_2\|^2 + a^2 b^2 \det(\mathbf{H}^\dagger \mathbf{H})} \right).$$

Next, we do the following steps:

Step 0) Set $\mathbf{Q}_s = \text{diag}[P_s/2, P_s/2]$.

Step 1) Calculate $\mathbf{Q}^e = \mathbf{GQ}_s \mathbf{G}^\dagger$.

$$\text{Set } a^{peak} = \sqrt{\frac{P_{r1}}{1+q_{11}}}, b^{peak} = \sqrt{\frac{P_{r2}}{1+q_{22}}}, \text{ and } \phi = \angle(\mathbf{h}_1^\dagger \mathbf{h}_2 q_{21}).$$

Step 2) Calculate $\max_a R(a, b^{peak})$ subject to $0 \leq a \leq a^{peak}$ by solving $\frac{\partial}{\partial a} R(a, b^{peak}) = 0$.

Calculate $\max_b R(a^{peak}, b)$ subject to $0 \leq b \leq b^{peak}$ by solving $\frac{\partial}{\partial b} R(a^{peak}, b) = 0$.

Between (a^{opt}, b^{peak}) and (a^{peak}, b^{opt}) , choose the one that results in a higher rate.

Step 3) Set $\mathbf{A} = \text{diag}[ae^{j\phi}, b]$ using the values obtained above.

Calculate \mathbf{K} and do eigenvalue decomposition to get \mathbf{U} and Λ such that

$$\mathbf{K} = \mathbf{I} + \mathbf{HAA}^\dagger \mathbf{H}^\dagger = \mathbf{U}^\dagger \Lambda \mathbf{U}.$$

$$\text{Calculate } \mathbf{F} = \Lambda^{-1/2} \mathbf{U} \mathbf{H} \mathbf{A} \mathbf{G}.$$

Step 4) Optimize \mathbf{Q}_s via singular value decomposition of \mathbf{F} and waterfilling [19].

Update \mathbf{Q}_s and calculate $R_{AF} = \log \det(\mathbf{I} + \mathbf{FQ}_s \mathbf{F}^\dagger)$.

Step 5) Terminate if R_{AF} already converged to a certain value.

Otherwise, go to Step 1.

Before closing the section, it is worth noting that full power transmission sometimes hurts. This is the case when one of the relays has received a signal with very low SNR so that transmission at full power degrades the received SNR at the destination. Fig. 9 shows examples. For each of the two different values of $\|\mathbf{g}_2\|$ shown in the figure, we run the algorithm steps from 0 to 4 just once, and draw the resulting $R(a^{peak}, b)$ versus b curves. For the curve in Fig. 9. (a), the maximum of the curve indicated by a circle happens at a point of b above b^{peak} indicated by vertical line. It means that the received SNR at the second relay is still high, and full power transmission helps. However, for the curve in Fig. 9. (b), as the maximum happens at a point less than b^{peak} , the received SNR at the second relay is too low for transmission at full power to be optimal.

VI. COMPRESS-AND-FORWARD (CF)

For CF, relays compress or quantize their received signals, re-encode, and forward them to the destination. At the destination, the decoder tries to decode the received signal to recover relay input signals, and finally decompress the relay signals to recover the information transmitted from the source. CF achievable rates were first derived by applying Wyner-Ziv source coding [15] to a classical one relay model in [3], their extension to multiple relay models in [4]. The derivation for a special case of the Gaussian parallel relay network with $N = M = 1$ can be found in [1] and [2]. The extension to our network model in terms of mutual information is straightforward as follows

$$\begin{aligned}
 R_{CF} &\leq I(\mathbf{x}_s; \hat{y}_{r1}, \hat{y}_{r2}) \\
 I(\hat{y}_{r1}; y_{r1} | \hat{y}_{r2}) &\leq I(x_{r1}; \mathbf{y}_d | x_{r2}) \\
 I(\hat{y}_{r2}; y_{r2} | \hat{y}_{r1}) &\leq I(x_{r2}; \mathbf{y}_d | x_{r1}) \\
 I(\hat{y}_{r1}, \hat{y}_{r2}; y_{r1}, y_{r2}) &\leq I(x_{r1}, x_{r2}; \mathbf{y}_d).
 \end{aligned} \tag{13}$$

By introducing quantization noise, we have a compressed version of relay signals given by

$$\begin{aligned}
 \hat{y}_{r1} &= y_{r1} + \sqrt{a} \hat{n}_{r1} \\
 \hat{y}_{r2} &= y_{r2} + \sqrt{b} \hat{n}_{r2}
 \end{aligned} \tag{14}$$

where $a, b > 0$. Assuming the quantization noise and all input signals are Gaussian distributed, we can evaluate the mutual information expressions to have the following result.

Theorem 7: With CF, the following rate is achievable in the vector Gaussian parallel relay network.

$$\begin{aligned}
 R_{CF} &= \max_{\mathbf{Q}_s, \mathbf{A}} \log \frac{\det(\mathbf{I} + \mathbf{A} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger)}{(1+a)(1+b)}, \\
 \text{subject to } & \log \frac{\det(\mathbf{I} + \mathbf{A} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger)}{a(1+b+\mathbf{g}_2\mathbf{Q}_s\mathbf{g}_2^\dagger)} \leq \log(1+P_{r1}\|\mathbf{h}_1\|^2), \\
 & \log \frac{\det(\mathbf{I} + \mathbf{A} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger)}{b(1+a+\mathbf{g}_1\mathbf{Q}_s\mathbf{g}_1^\dagger)} \leq \log(1+P_{r2}\|\mathbf{h}_2\|^2), \\
 & \log \frac{\det(\mathbf{I} + \mathbf{A} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger)}{ab} \leq \log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger), \\
 & \text{tr}(\mathbf{Q}_s) \leq P_s, \\
 & a, b \geq 0
 \end{aligned} \tag{15}$$

where $\mathbf{Q}_s = \mathbb{E}[\mathbf{x}_s\mathbf{x}_s^\dagger]$, $\mathbf{Q}_r = \mathbb{E}[\mathbf{x}_r\mathbf{x}_r^\dagger] = \text{diag}[P_{r1}, P_{r2}]$, and $\mathbf{A} = \text{diag}[a, b]$.

Proof: It is straightforward to show the theorem result by evaluating the mutual information expressions with the assumption that the input distributions are circularly symmetric complex Gaussian. \blacksquare

Similar to the AF scheme, the CF achievable rate becomes close to the upper bound as the relay power goes to infinity. We get the following result for the CF achievable rate.

Theorem 8 (Asymptotic Optimality of CF): In the vector Gaussian parallel relay network, regardless of the rank of \mathbf{H} , CF is asymptotically optimal in the high relay power limit in the sense that

$$\lim_{P_{r1}=P_{r2} \rightarrow \infty} (\max_{\mathbf{Q}_s} \log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger) - R_{CF}) = 0.$$

Proof: As P_{r1} and P_{r2} go to infinity, the optimization of R_{CF} becomes unconstrained. The objective is maximum when $a = b = 0$. Thus,

$$\max_{\mathbf{Q}_s, \mathbf{A}} \log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger(\mathbf{I} + \mathbf{A})^{-1}) \rightarrow \max_{\mathbf{Q}_s} \log \det(\mathbf{I} + \mathbf{G}\mathbf{Q}_s\mathbf{G}^\dagger). \tag{16}$$

\blacksquare

VII. NUMERICAL RESULTS

In this section, we consider a few numerical examples to compare achievable rates by different schemes and the upper bound derived throughout the paper. Let us pick three symmetric matrices for \mathbf{G} (or \mathbf{H}):

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0.9285 & 0.3714 \\ 0.3714 & 0.9285 \end{bmatrix}, \text{ and } \begin{bmatrix} 0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}.$$

The angles between channel row (or column) vectors are 90° , 46.3972° , and 0° , respectively. Fig. 10 shows the results of six different combinations of the first and second hop channel matrices where the achievable rates by DF, AF and CF, and the upper bound are plotted.

As we have investigated in Section IV, below a certain level of the relay-to-destination link SNR, DF achievable rate meets the upper bound. The threshold point at which DF starts achieving the capacity can be calculated from the DF optimality condition. In Fig. 10. (a) and (b), we can see that when the first hop channel vectors are orthogonal, DF always performs better than AF and CF, and achieves the capacity in the high relay power regime. In contrast, when the channel vectors are not orthogonal as in Fig. 10. (c), (d), (e) and (f), DF achievable rates are bounded away from the upper bound in the high relay power regime.

When the second hop channel matrix is full rank as in Fig. 10. (a), (c) and (e), AF is shown to asymptotically achieve the capacity in the high relay power limit. In Fig. 10. (b) and (d), AF achievable rate stays away from the capacity even in the high relay power limit since the second hop channel is rank-deficient. In Fig. 10. (f), again, AF becomes asymptotically optimal even though the second hop channel vectors are not full rank. In the case, as the first hop channel is already rank-deficient, there is no additional penalty by the rank-deficient second hop. CF seems to be advantageous over AF in the sense that it asymptotically achieve the capacity in the high relay power limit regardless of the rank of the second hop channel.

VIII. CONCLUSION

Throughout the paper, we have shown how much rate is achievable by DF, AF or CF, and when the achievable rates meet the upper bound. The relative advantage of each scheme varies depending not only on which hop is the bottleneck but also on the ranks of the first and second hop channel matrices. The effect of the channel rank is newly explained in our work.

For the DF relaying, we used a combination of a MISO broadcast scheme and a SIMO multiple access scheme, with which a few interesting characteristics of the SIMO MAC are investigated. It is shown that DF achieves the capacity in the low relay power regime.

Earlier results for AF and CF were extended to explain our vector Gaussian network and to compare their achievable rates to that of DF. AF was shown to achieve close-to-capacity rate in the high relay power regime when the second hop channel matrix is full rank while CF similarly achieves the asymptotic capacity regardless of the channel rank.

APPENDIX I PROOF OF LEMMA 1

First, we shall find the optimal angle of ρ . By differentiating $\det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) = \det(\mathbf{I} + \mathbf{Q}_r\mathbf{H}^\dagger\mathbf{H})$ with respect to θ and setting it to zero, we get

$$e^{j\theta}\mathbf{h}_2^\dagger\mathbf{h}_1 = e^{-j\theta}\mathbf{h}_1^\dagger\mathbf{h}_2.$$

Since the left hand side is the conjugate of the right hand side, they both should be real, which means the optimal θ needs to satisfy

$$\theta^{opt} = -\angle(\mathbf{h}_2^\dagger\mathbf{h}_1) = \angle(\mathbf{h}_1^\dagger\mathbf{h}_2).$$

Using this optial angle, we can solve the following convex optimization problem to find the optimal $|\rho|$:

$$\max_{|\rho|} \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger)$$

$$\text{subject to } 0 \leq |\rho| \leq 1$$

where its Lagrangian function is given by

$$L(|\rho|, \lambda) = -\det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger) + \lambda_1(|\rho| - 1) + \lambda_2(-|\rho|).$$

The Karush-Kuhn-Tucker (KKT) condition is given by

$$\nabla L(|\rho|, \lambda) = 2|\rho|\sqrt{P_{r1}P_{r2}}\det(\mathbf{H}^\dagger\mathbf{H}) - 2|\mathbf{h}_1^\dagger\mathbf{h}_2| + \lambda_1 - \lambda_2 = 0.$$

Solving this for $|\rho|$ gives

$$|\rho| = \frac{|\mathbf{h}_1^\dagger\mathbf{h}_2| + (\lambda_2 - \lambda_1)/2}{\sqrt{P_{r1}P_{r2}}\det(\mathbf{H}^\dagger\mathbf{H})}.$$

From complementary slackness, it must be satisfied that $\lambda_1(|\rho| - 1) = 0$ and $\lambda_2|\rho| = 0$. Thus, if $0 < |\rho| < 1$, then the optimal solution would be $\lambda_1 = 0$, $\lambda_2 = 0$, and

$$|\rho| = \frac{|\mathbf{h}_1^\dagger \mathbf{h}_2|}{\sqrt{P_{r1}P_{r2}} \det(\mathbf{H}^\dagger \mathbf{H})}.$$

Likewise, we also have the following two sets of solutions,

$$|\rho| = 0, \lambda_1 = 0, \lambda_2 = 2|\mathbf{h}_1^\dagger \mathbf{h}_2|$$

$$|\rho| = 1, \lambda_1 = 2\sqrt{P_{r1}P_{r2}} \det(\mathbf{H}^\dagger \mathbf{H}) - 2|\mathbf{h}_1^\dagger \mathbf{h}_2|, \lambda_2 = 0.$$

In other words, the function $\log \det(\mathbf{I} + \mathbf{H}\mathbf{Q}_r\mathbf{H}^\dagger)$ is a quadratic and concave function of $|\rho|$ with its maximum at $|\rho|' = \frac{|\mathbf{h}_1^\dagger \mathbf{h}_2|}{\sqrt{P_{r1}P_{r2}} \det(\mathbf{H}^\dagger \mathbf{H})} \geq 0$ without constraints. If $|\rho|' \leq 1$, the constraint is inactive so $|\rho|'$ maximizes the objective function. If $|\rho|' > 1$, it violates the constraint, and the objective function has its maximum at the boundary of the feasible set $|\rho| = 1$.

REFERENCES

- [1] B. E. Schein, *Distributed coordination in network information theory*, Ph.D. dissertation Cambridge, MA: MIT, Oct. 2001.
- [2] F. Xue and S. Sandhu, "Cooperation in a half-duplex Gaussian diamond relay channel," *IEEE Trans. Inf. Theory*, vol. 53, no. 10, pp. 3806-3814, Oct. 2007.
- [3] T. M. Cover and A. A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 5, pp. 572-584, Sep. 1979.
- [4] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037-3063, Sep. 2005.
- [5] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062-3080, Dec. 2004.
- [6] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 29-43, Jan. 2005.
- [7] H. Bölcseki, R. U. Nabar, Ö. Oyman, and A. J. Paulraj, "Capacity scaling laws in MIMO relay networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 6, pp. 1433-1444, Jun. 2006.
- [8] Y. Liang and V. V. Veeravalli, "Cooperative relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, no. 3, pp. 900-928, Mar. 2007.
- [9] A. Wittneben and B. Rankow, "Impact of cooperative relays on the capacity of rank-deficient MIMO channels," in *Proc. 12th IST Summit on Mob. and Wirel. Comm.*, Aveiro, Portugal, June 2003, pp. 421-425.
- [10] N. Jindal and A. Goldsmith, "Optimal power allocation for parallel broadcast channels with independent and common information," in *Proc. IEEE Int. Symp. Information Theory*, Chicago, IL, June 2004, p. 215.
- [11] H. Weingarten, Y. Steinberg, and S. Shamai, "On the capacity region of the multi-antenna broadcast channel with common messages," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, July 2006, p. 2195-2199.

- [12] D. Wajcer, S. Shamai, and A. Wiesel, "On superposition coding and beamforming for the multi-antenna Gaussian broadcast channels," *Information Theory and Applications Workshop*, UCSD, 2006.
- [13] M. H. M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439-441, May 1983.
- [14] D. Slepian and J. K. Wolf, "A coding theorem for multiple access channels with correlated sources," *Bell Syst. Tech. J.*, vol. 52, pp. 1037-1076, Sept. 1973.
- [15] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the receiver," *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 1-11, Jan. 1976.
- [16] F. M. J. Willems, *Informationtheoretical results for the discrete memoryless multiple access channel*, Doctor in de Wetenschappen Proefschrift dissertation, Katholieke Universiteit Leuven, Leuven, Belgium, Oct. 1982.
- [17] V. V. Prelov and E. C. van der Meulen, "Asymptotic expansion for the capacity region of the multiple-access channel with common information and almost Gaussian noise," in *Proc. IEEE Int. Symp. Information Theory*, Budapest, Hungary, June 1991, p. 300.
- [18] N. Liu and S. Ulukus, "Capacity region and optimum power control strategies for fading Gaussian multiple access channels with common data," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1815-1826, Oct. 2006.
- [19] İ. E. Telatar, "Capacity of multi-antenna Gaussian channels," *Europ. Trans. Telecommun.*, vol. 10, pp. 585-596, Nov.-Dec. 1999.
- [20] D. N. C. Tse and P. Viswanath, *Fundamentals of wireless communication*, Cambridge University Press, May 2005.
- [21] T. Cover and J. Thomas, *Elements of information theory*, New York: Wiley, 1991.
- [22] A. S. Avestimehr, S. N. Diggavi and D. N. C. Tse, "Approximate capacity of Gaussian relay networks," in *Proc. IEEE Int. Symp. Information Theory*, Toronto, Canada, July 2008, p. 474-478.
- [23] Y. R. Fan, C. Wang, H. V. Poor and J. S. Thompson, "Cooperative multiplexing: toward higher spectral efficiency in multiple-antenna relay networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 3909-3926, Sep. 2009.
- [24] S. S. C. Rezaei, S. O. Gharan and A. K. Khandani, "A new achievable rate for the Gaussian parallel relay channel," in *Proc. IEEE Int. Symp. Information Theory*, Seoul, Korea, June-July 2009, p. 194-198.

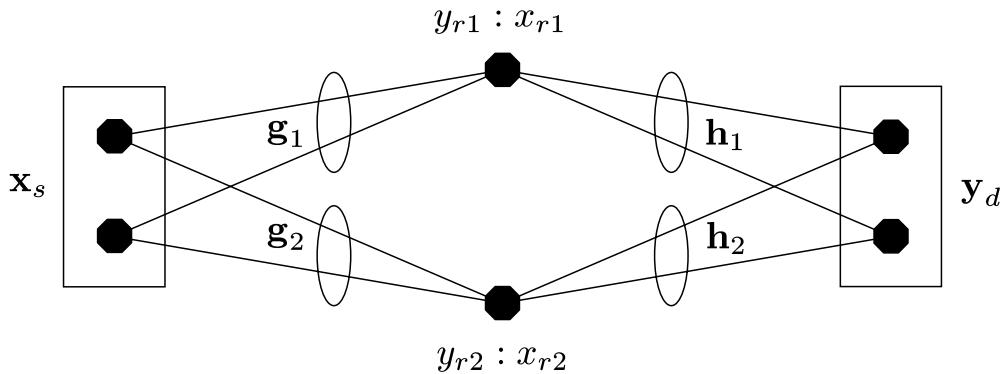


Fig. 1. System model

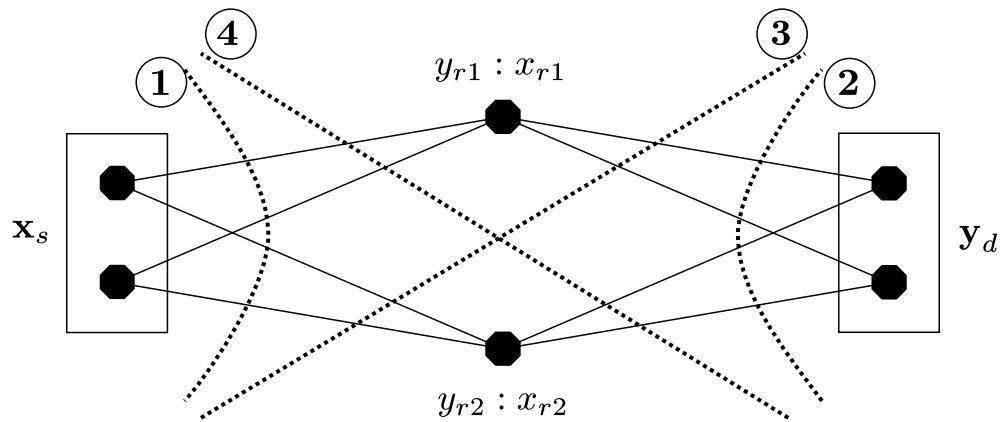
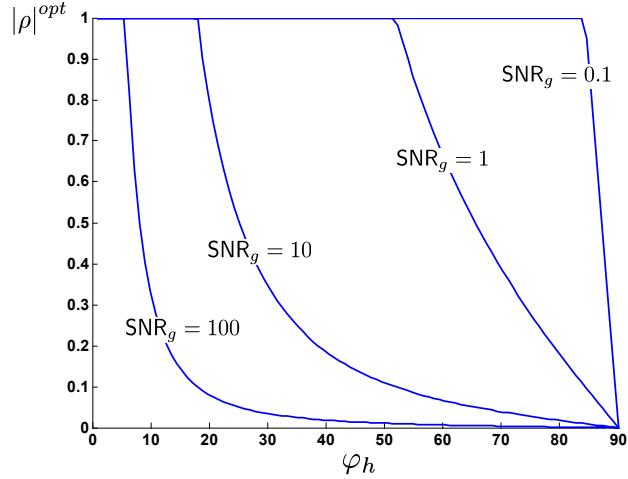
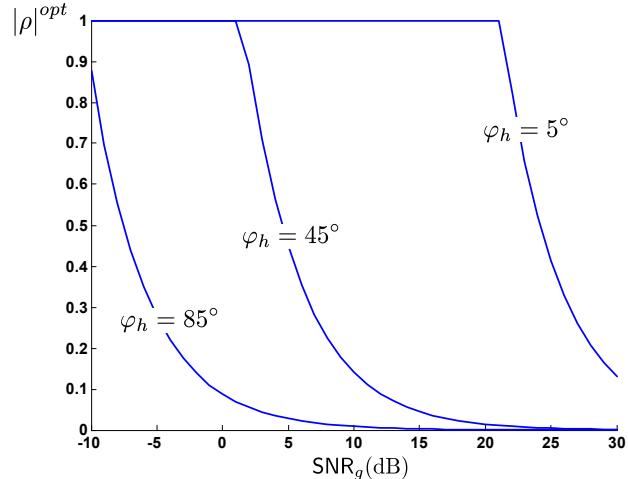


Fig. 2. Cut-set upper bounds

(a) Optimal correlation versus φ_h (b) Optimal correlation versus SNR_g Fig. 3. Optimal correlation as a function of φ_h and SNR_g

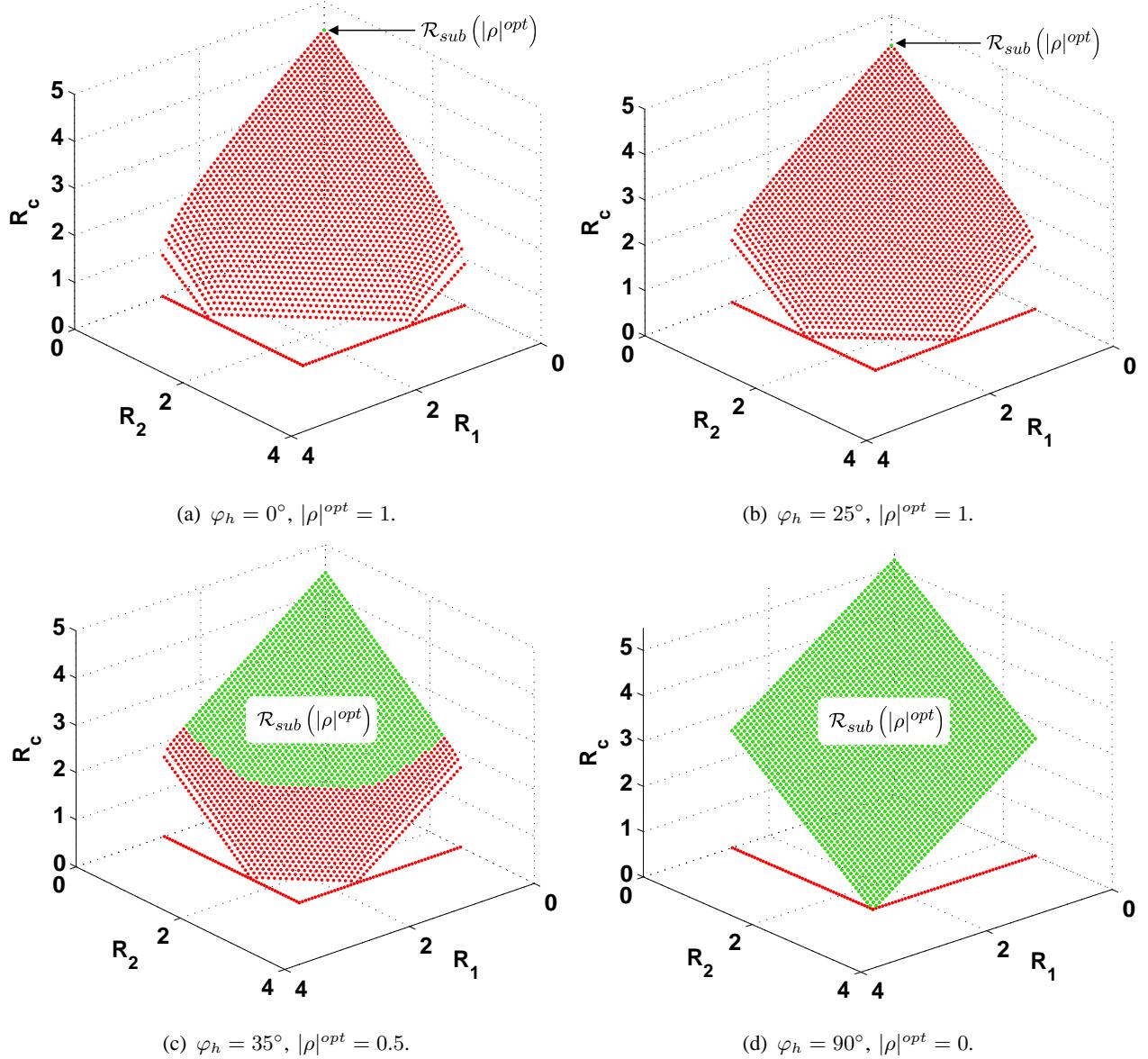


Fig. 4. SIMO MAC achievable regions with $P_{r1} = P_{r2} = 5$ and $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$

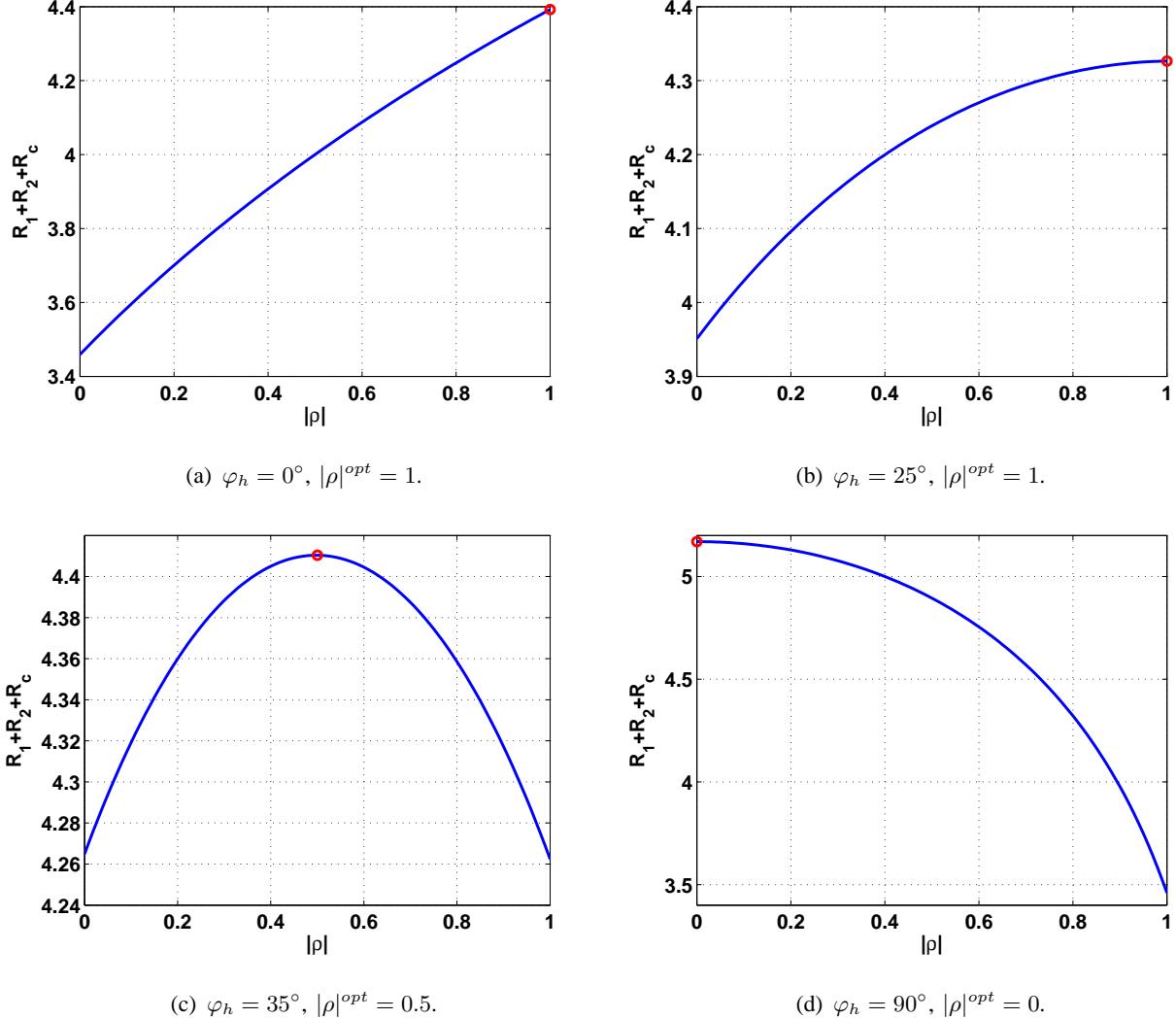
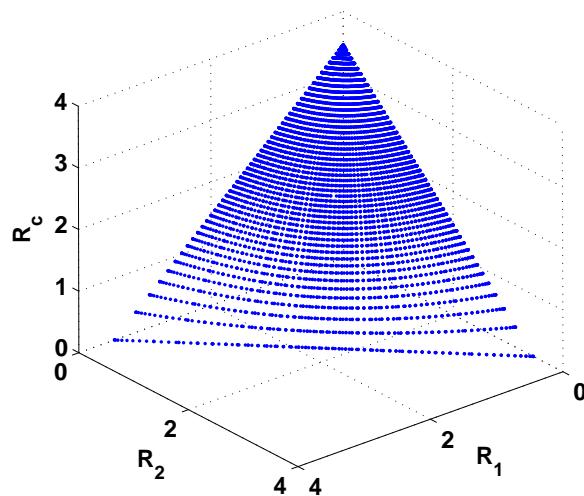
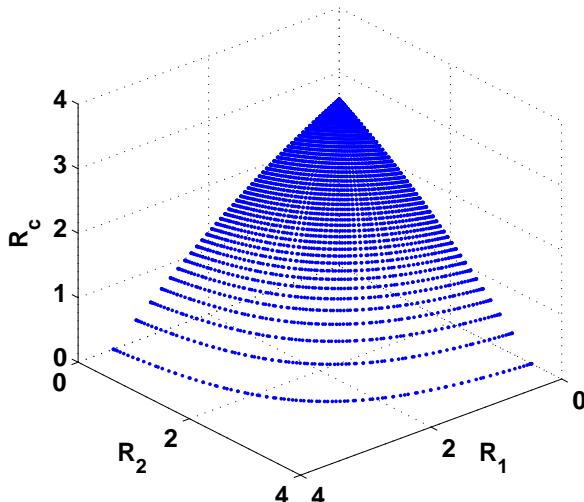
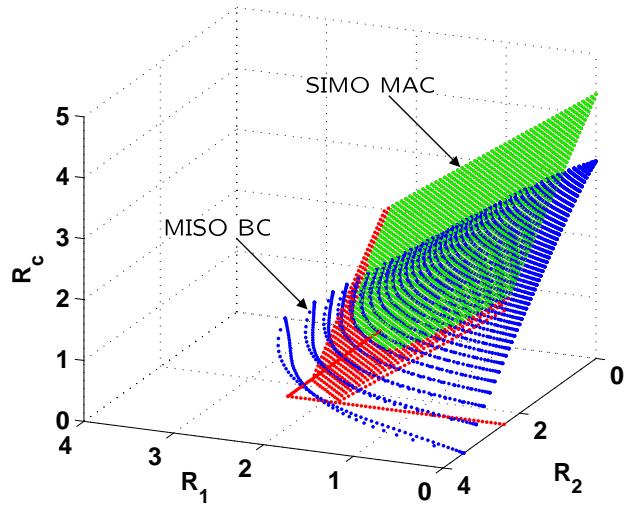
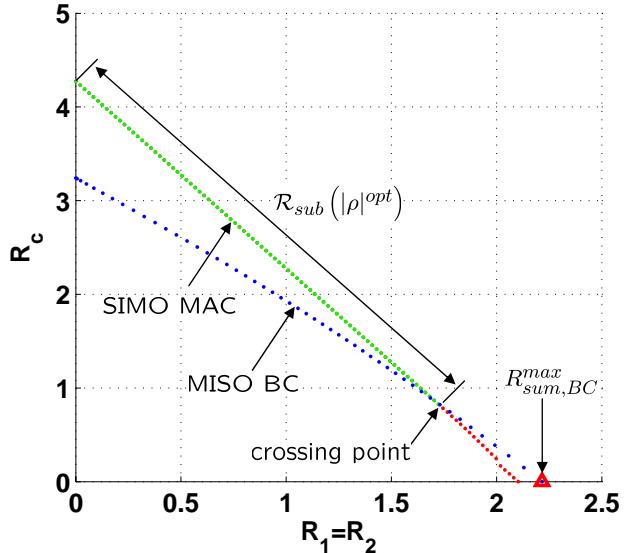


Fig. 5. Sum-rate versus $|\rho|$ with $P_{r1} = P_{r2} = 5$ and $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$

(a) Parallel channel vectors: $\varphi_g = 0^\circ$ (b) Orthogonal channel vectors: $\varphi_g = 90^\circ$ Fig. 6. MISO BC achievable regions with $P_s = 10$ and $\|\mathbf{g}_1\| = \|\mathbf{g}_2\| = 1$



(a) Three-dimensional region

(b) Cross section along the surface $R_1 = R_2$ Fig. 7. DF achievable region: $P_s = 10$, $P_{r1} = P_{r2} = 4.17$, $\varphi_g = \varphi_h = 46.3942^\circ$, $|\rho|^{opt} = 0.31$ and $R_c^{th} = 0.79$.

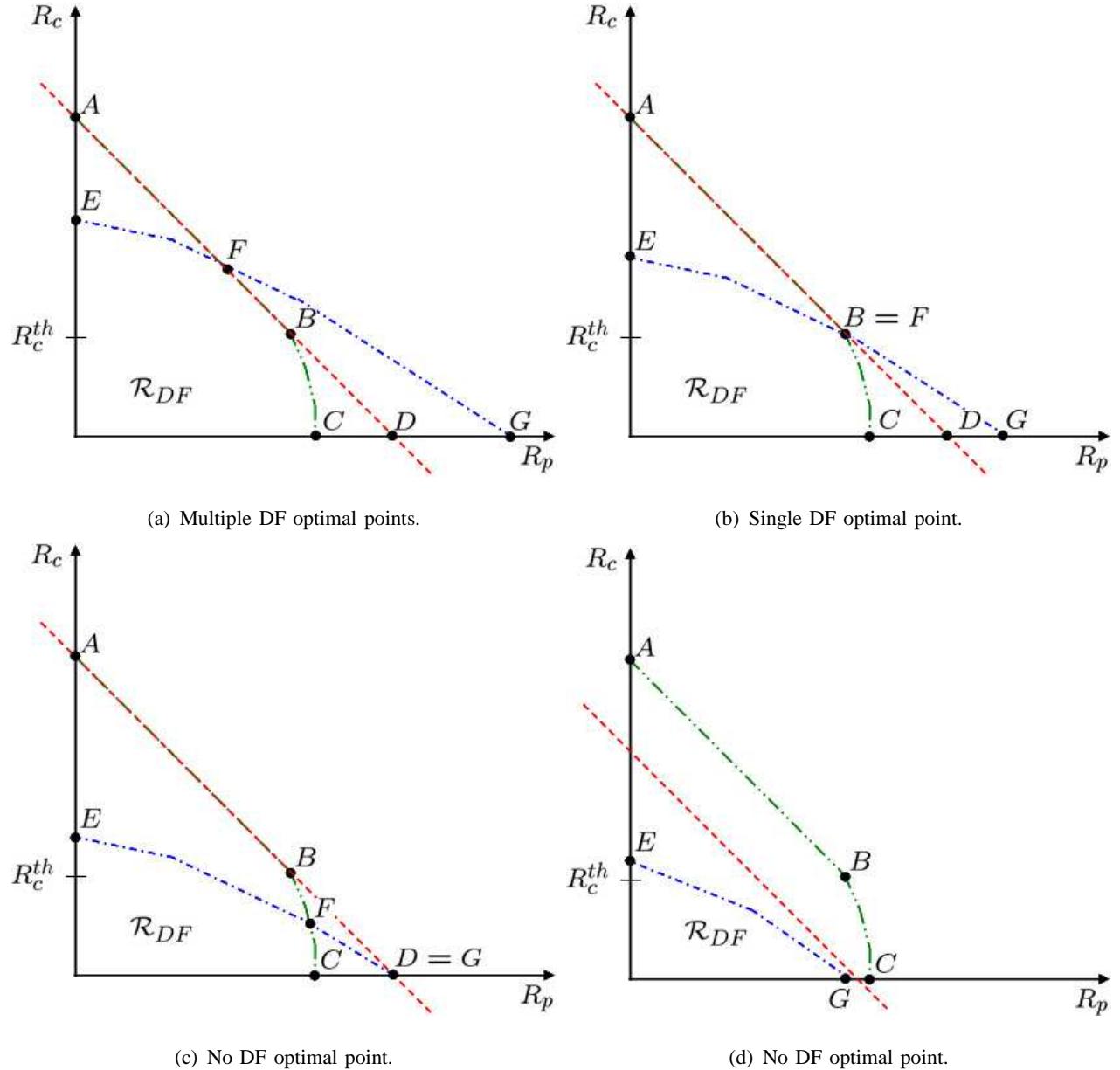
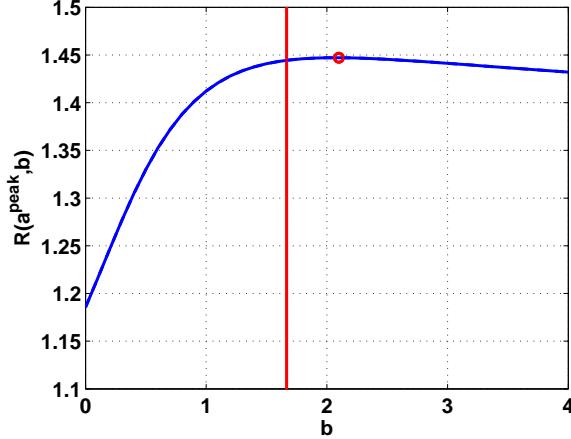
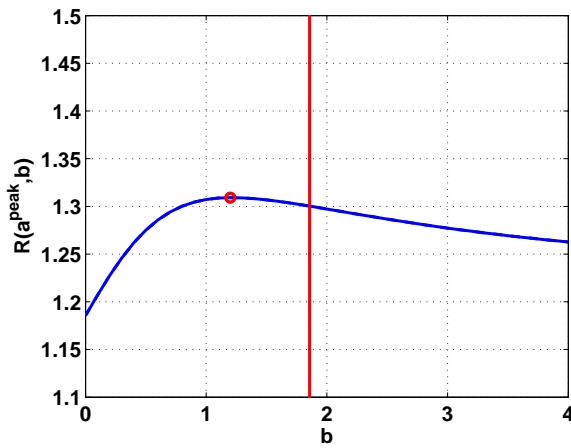


Fig. 8. Symmetric channels.

(a) $\|\mathbf{g}_1\| = 1$, $\|\mathbf{g}_2\| = 0.16$ and $b^{peak} = 1.67$ (b) $\|\mathbf{g}_1\| = 1$, $\|\mathbf{g}_2\| = 0.09$ and $b^{peak} = 1.85$ Fig. 9. $R(a^{peak}, b)$ versus b with $P_s = 10$, $P_{r1} = P_{r2} = 5$, $\|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$ and $\varphi_g = \varphi_h = 46.40^\circ$

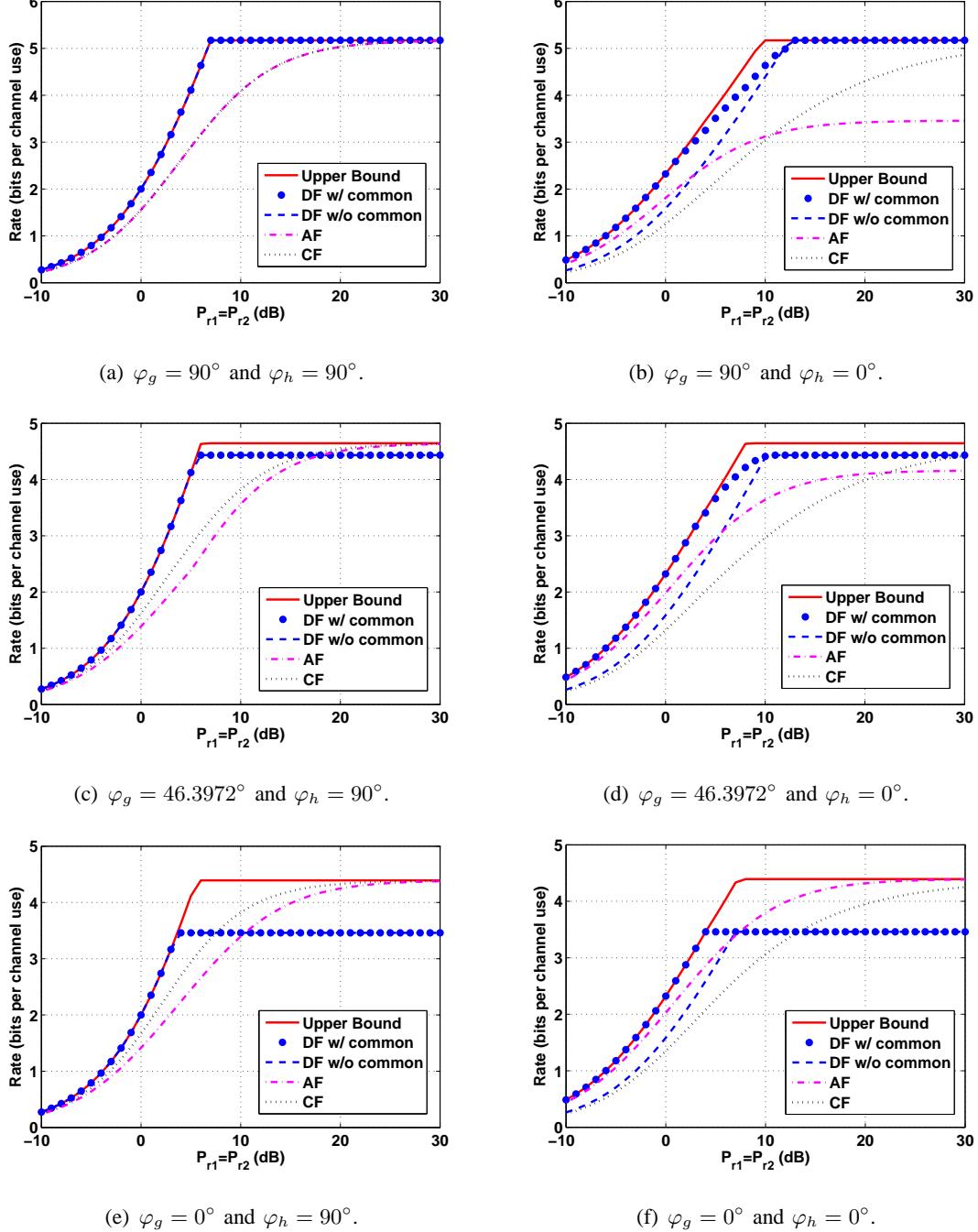


Fig. 10. Upper bound and achievable rates versus relay power. Note that $P_s = 10$ (dB) and $\|\mathbf{g}_1\| = \|\mathbf{g}_2\| = \|\mathbf{h}_1\| = \|\mathbf{h}_2\| = 1$.